

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2001, 09:33:09 ; Search time 15.54 Seconds  
(without alignments)  
1717.187 Million cell updates/sec

Title: US-09-360-934A-3  
Perfect score: 1296  
Sequence: 1 MEIQTHRKINRPLVSLV.....HNLISNIGHFASNLGRYSF 1296

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 200 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1296	100.0	1296	3	US-08-470-260-3
2	1296	100.0	1296	3	US-08-471-491-3
3	1296	100.0	1296	4	US-08-466-662-3
4	99	7.6	1287	1	US-08-200-232-2
5	99	7.6	1287	5	PCT-US95-02219-2
6	99	7.6	1287	5	PCT-US95-02219A-2
7	23	1.8	23	3	US-08-295-643-2
8	23	1.8	23	3	US-08-473-265-1
9	23	1.8	23	3	US-08-284-747-1
10	20	1.5	513	1	US-08-200-232-4
11	20	1.5	513	5	PCT-US95-02219-4
12	20	1.5	513	5	PCT-US95-02219A-4
13	14	1.1	17	2	US-08-295-643-15
14	10	0.8	21	2	US-08-295-643-17
15	8	0.6	9	2	US-08-295-643-16
16	7	0.5	13	5	PCT-US95-04121-43
17	7	0.5	13	5	PCT-US95-04121-46
18	7	0.5	16	2	US-08-337-646A-24
19	7	0.5	16	4	US-08-927-326-24
20	7	0.5	17	2	US-08-207-481-5
21	7	0.5	17	2	US-08-640-344-10
22	7	0.5	17	2	US-08-640-344-11
23	7	0.5	17	5	PCT-US95-02689-5
24	7	0.5	20	1	US-08-418-893D-8
25	7	0.5	20	4	US-09-007-905-59
26	7	0.5	20	4	US-09-007-905-64
27	7	0.5	22	4	US-09-166-028-7
28	7	0.5	7	7	Sequence 3, Appli
29	7	0.5	7	7	Sequence 3, Appli
30	7	0.5	7	7	Sequence 3, Appli
31	7	0.5	7	7	Sequence 2, Appli
32	7	0.5	7	7	Sequence 2, Appli
33	7	0.5	7	7	Sequence 2, Appli
34	7	0.5	7	7	Sequence 2, Appli
35	7	0.5	7	7	Sequence 1, Appli
36	7	0.5	7	7	Sequence 1, Appli
37	7	0.5	7	7	Sequence 4, Appli
38	7	0.5	7	7	Sequence 4, Appli
39	7	0.5	7	7	Sequence 4, Appli
40	7	0.5	7	7	Sequence 15, Appli
41	7	0.5	7	7	Sequence 17, Appli
42	7	0.5	7	7	Sequence 16, Appli
43	7	0.5	7	7	Sequence 43, Appli
44	7	0.5	7	7	Sequence 46, Appli
45	7	0.5	7	7	Sequence 24, Appli
46	7	0.5	7	7	Sequence 5, Appli
47	7	0.5	7	7	Sequence 10, Appli
48	7	0.5	7	7	Sequence 11, Appli
49	7	0.5	7	7	Sequence 5, Appli
50	7	0.5	7	7	Sequence 8, Appli
51	7	0.5	7	7	Sequence 59, Appli
52	7	0.5	7	7	Sequence 64, Appli
53	7	0.5	7	7	Sequence 7, Appli
54	7	0.5	7	7	Sequence 3, Appli
55	7	0.5	7	7	Sequence 3, Appli
56	7	0.5	7	7	Sequence 38, Appli
57	7	0.5	7	7	Sequence 13, Appli
58	7	0.5	7	7	Sequence 19, Appli
59	7	0.5	7	7	Sequence 13, Appli
60	7	0.5	7	7	Sequence 62, Appli
61	7	0.5	7	7	Sequence 6, Appli
62	7	0.5	7	7	Sequence 17, Appli
63	7	0.5	7	7	Sequence 18, Appli
64	7	0.5	7	7	Sequence 21, Appli
65	7	0.5	7	7	Sequence 51, Appli
66	7	0.5	7	7	Sequence 10, Appli
67	7	0.5	7	7	Sequence 10, Appli
68	7	0.5	7	7	Sequence 12, Appli
69	7	0.5	7	7	Sequence 20, Appli
70	7	0.5	7	7	Sequence 12, Appli
71	7	0.5	7	7	Sequence 12, Appli
72	7	0.5	7	7	Sequence 4, Appli
73	7	0.5	7	7	Sequence 8, Appli
74	7	0.5	7	7	Sequence 10, Appli
75	7	0.5	7	7	Sequence 12, Appli
76	7	0.5	7	7	Sequence 13, Appli
77	7	0.5	7	7	Sequence 17, Appli
78	7	0.5	7	7	Sequence 5, Appli
79	7	0.5	7	7	Sequence 20, Appli
80	7	0.5	7	7	Sequence 20, Appli
81	7	0.5	7	7	Sequence 15, Appli
82	7	0.5	7	7	Sequence 17, Appli
83	7	0.5	7	7	Sequence 10, Appli
84	7	0.5	7	7	Sequence 24, Appli
85	7	0.5	7	7	Sequence 30, Appli
86	7	0.5	7	7	Sequence 27, Appli
87	7	0.5	7	7	Sequence 109, App
88	7	0.5	7	7	Sequence 109, App
89	7	0.5	7	7	Sequence 114, App
90	7	0.5	7	7	Sequence 114, App
91	7	0.5	7	7	Sequence 11, Appli
92	7	0.5	7	7	Sequence 18, Appli
93	7	0.5	7	7	Sequence 26, Appli
94	7	0.5	7	7	Sequence 26, Appli
95	7	0.5	7	7	Sequence 26, Appli
96	7	0.5	7	7	Sequence 26, Appli
97	7	0.5	7	7	Sequence 26, Appli
98	7	0.5	7	7	Sequence 26, Appli
99	7	0.5	7	7	Sequence 26, Appli
100	7	0.5	7	7	Sequence 26, Appli

Sequence 1, Appli  
Sequence 36, Appli  
Sequence 40, Appli  
Sequence 45, Appli  
Sequence 52, Appli  
Sequence 50, Appli  
Sequence 4, Appli  
Patent No. 5514582  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 38, Appli  
Sequence 13, Appli  
Sequence 19, Appli  
Sequence 13, Appli  
Sequence 12, Appli  
Sequence 62, Appli  
Sequence 6, Appli  
Sequence 17, Appli  
Sequence 18, Appli  
Sequence 21, Appli  
Sequence 22, Appli  
Sequence 9, Appli  
Sequence 21, Appli  
Sequence 51, Appli  
Sequence 10, Appli  
Sequence 10, Appli  
Sequence 12, Appli  
Sequence 20, Appli  
Sequence 12, Appli  
Sequence 12, Appli  
Sequence 4, Appli  
Sequence 21, Appli  
Sequence 51, Appli  
Sequence 15, Appli  
Sequence 17, Appli  
Sequence 10, Appli  
Sequence 12, Appli  
Sequence 16, Appli  
Sequence 10, Appli  
Sequence 10, Appli  
Sequence 12, Appli  
Sequence 4, Appli  
Sequence 21, Appli  
Sequence 8, Appli  
Sequence 10, Appli  
Sequence 12, Appli  
Sequence 13, Appli  
Sequence 17, Appli  
Sequence 5, Appli  
Sequence 20, Appli  
Patent No. 5459251  
Patent No. 5506344  
Sequence 24, Appli  
Sequence 26, Appli  
Sequence 28, Appli  
Sequence 30, Appli  
Sequence 27, Appli  
Sequence 109, App  
Sequence 109, App  
Sequence 114, App  
Sequence 114, App  
Sequence 11, Appli  
Sequence 39, Appli  
Sequence 41, Appli  
Sequence 1, Appli  
Sequence 18, Appli  
Sequence 26, Appli  
Sequence 26, Appli  
Sequence 26, Appli





LENGTH: 1296 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-470-260-3

Query Match 100.0%; Score 1296; DB 3; Length 1296;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIQOTHRKINRPLVSLALVGVLSITPQSHAAFTTIIIPAIIVGGIATGTAAGTAVGSL 60  
DB 1 MEIQOTHRKINRPLVSLALVGVLSITPQSHAAFTTIIIPAIIVGGIATGTAAGTAVGSL 60

QY 61 LSWGLKQAEAEANKTPDKPKVWRIOAGKGFNEFPNKEYDLYRSLSSSKIDGGWDMGNAAR 120  
DB 61 LSWGLKQAEAEANKTPDKPKVWRIOAGKGFNEFPNKEYDLYRSLSSSKIDGGWDMGNAAR 120

QY 121 HYWKGGQONKLEVDKDAVGTYTISGLRNFETGGDLVNMOKATLRLQGFNGNSFTSYKD 180  
DB 121 HYWKGGQONKLEVDKDAVGTYTISGLRNFETGGDLVNMOKATLRLQGFNGNSFTSYKD 180

QY 181 SADRTTRVDFAKNKISIDNFEINNRVSGAGKASSVTLTQASEGITSKNAEISLYD 240  
DB 181 SADRTTRVDFAKNKISIDNFEINNRVSGAGKASSVTLTQASEGITSKNAEISLYD 240

QY 241 GATLNLASSSVKLMGNVWMLQYVAYLAPSYSTINTSKVTGEVNFHNTVGDKNAAQA 300  
DB 241 GATLNLASSSVKLMGNVWMLQYVAYLAPSYSTINTSKVTGEVNFHNTVGDKNAAQA 300

QY 301 GIANKKNTIGTDLQWAGNLIIAPPGGYKDKPNPTPSQSGAKNDKSNESAKNDQESS 360  
DB 301 GIANKKNTIGTDLQWAGNLIIAPPGGYKDKPNPTPSQSGAKNDKSNESAKNDQESS 360

QY 361 QNNSQTQVNPNSAQTEVQTVIDGPFAGGKDTVNIINRINTNADGTIRVGFKASL 420  
DB 361 QNNSQTQVNPNSAQTEVQTVIDGPFAGGKDTVNIINRINTNADGTIRVGFKASL 420

QY 421 TTNAHLHGKGGVNLNSOAGRSILVENLTGNTITVDGPRVNVNQGVALAGSANSFEF 480  
DB 421 TTNAHLHGKGGVNLNSOAGRSILVENLTGNTITVDGPRVNVNQGVALAGSANSFEF 480

QY 481 KAGDTKNGTATFNNDISLGRFVNLKVDATFANKGIDTNGNGFNTLDFSGVTDKVNINK 540  
DB 481 KAGDTKNGTATFNNDISLGRFVNLKVDATFANKGIDTNGNGFNTLDFSGVTDKVNINK 540

QY 541 LITASTNVAVKFNELIVKTNGLSVGEYTHFSEDIGSQSRINTVRLTGTGTRSLFSGV 600  
DB 541 LITASTNVAVKFNELIVKTNGLSVGEYTHFSEDIGSQSRINTVRLTGTGTRSLFSGV 600

QY 601 KFKGGEKLIVDEFYSPWNYEDARNIKNVEITNKLAFQPGSPWGTSKLMPFNLLTQNA 660  
DB 601 KFKGGEKLIVDEFYSPWNYEDARNIKNVEITNKLAFQPGSPWGTSKLMPFNLLTQNA 660

QY 661 VMDYSQSNLTIOGDFINNOQTINYLVRGKQVATLSVGNAAAMFNNDIDSATGYFKPLI 720  
DB 661 VMDYSQSNLTIOGDFINNOQTINYLVRGKQVATLSVGNAAAMFNNDIDSATGYFKPLI 720

QY 721 KINSAQDLIKNTEHVLKAKIIGYCNVSTGTNGISNVNLEQFERALYNNNRMDTCV 780  
DB 721 KINSAQDLIKNTEHVLKAKIIGYCNVSTGTNGISNVNLEQFERALYNNNRMDTCV 780

QY 781 VRNTDDIKACGMAIGDQSWNNPNKYLIKWKNTGISTANGSKISVYILGNSTPTE 840  
DB 781 VRNTDDIKACGMAIGDQSWNNPNKYLIKWKNTGISTANGSKISVYILGNSTPTE 840

QY 841 NGGNTNLTPTNTTSNARSANALAQNAFAQPSATPNLVAINHDFGTIESVFELANRSK 900  
DB 841 NGGNTNLTPTNTTSNARSANALAQNAFAQPSATPNLVAINHDFGTIESVFELANRSK 900

QY 901 DIDTLYANSAGQGRDLLQTLIDSHDAGYARKMIDATSANETIKQLTATTTLNIIASLE 960

DB 901 DIDTLYANSAGQGRDLLQTLIDSHDAGYARKMIDATSANETIKQLTATTTLNIIASLE 960  
QY 961 HKTSGLOTLSLSNAMILNSRLVNLRRHTNHIDSFARLQALKDOKFASLESAAEVLQF 1020  
DB 961 HKTSGLOTLSLSNAMILNSRLVNLRRHTNHIDSFARLQALKDOKFASLESAAEVLQF 1020

QY 1021 APKYKPTNVWANAIGCTSLNNGSNASLYGTSAGVDAYLNGQVEAIVGGFGSYGYSFNN 1080  
DB 1021 APKYKPTNVWANAIGCTSLNNGSNASLYGTSAGVDAYLNGQVEAIVGGFGSYGYSFNN 1080

QY 1081 RANLSNGANNNTFGYYSRIFANQHEFFDEAQAQALGSDOSSLNFKSALLQDLNQSHYLA 1140  
DB 1081 RANLSNGANNNTFGYYSRIFANQHEFFDEAQAQALGSDOSSLNFKSALLQDLNQSHYLA 1140

QY 1141 YSAATRASGYDYFAFPFNALVLKPSVGVSYNHLGTSNFKSNSTNOVALKNGSSOHLFNA 1200  
DB 1141 YSAATRASGYDYFAFPFNALVLKPSVGVSYNHLGTSNFKSNSTNOVALKNGSSOHLFNA 1200

QY 1201 SANVEARYYYGDTSYFYMNAGVLQEFARHVGSNNAASLNTFKVNAARNPLNTHARVAMGE 1260  
DB 1201 SANVEARYYYGDTSYFYMNAGVLQEFARHVGSNNAASLNTFKVNAARNPLNTHARVAMGE 1260

QY 1261 LKLAKEVFLNLGVVYLHNLISNIGHFASNLGMRYSF 1296  
DB 1261 LKLAKEVFLNLGVVYLHNLISNIGHFASNLGMRYSF 1296

RESULT 2

US-08-471-491-3  
; Sequence 3, Application US/08471491B  
; Patent No. 6090611  
; GENERAL INFORMATION:  
; APPLICANT: Covacci, Antonello  
; APPLICANT: Bugnoli, Massimo  
; APPLICANT: Telford, John  
; APPLICANT: Macchia, Giovanni  
; APPLICANT: Rappulli, Rino  
; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And  
; FILE REFERENCE: CHIR004  
; CURRENT APPLICATION NUMBER: US/08/471,491B  
; CURRENT FILING DATE: 1995-06-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1296  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-08-471-491-3

Query Match 100.0%; Score 1296; DB 3; Length 1296;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIQOTHRKINRPLVSLALVGVLSITPQSHAAFTTIIIPAIIVGGIATGTAAGTAVGSL 60  
DB 1 MEIQOTHRKINRPLVSLALVGVLSITPQSHAAFTTIIIPAIIVGGIATGTAAGTAVGSL 60

QY 61 LSWGLKQAEAEANKTPDKPKVWRIOAGKGFNEFPNKEYDLYRSLSSSKIDGGWDMGNAAR 120  
DB 61 LSWGLKQAEAEANKTPDKPKVWRIOAGKGFNEFPNKEYDLYRSLSSSKIDGGWDMGNAAR 120

QY 121 HYWKGGQONKLEVDKDAVGTYTISGLRNFETGGDLVNMOKATLRLQGFNGNSFTSYKD 180  
DB 121 HYWKGGQONKLEVDKDAVGTYTISGLRNFETGGDLVNMOKATLRLQGFNGNSFTSYKD 180

QY 181 SADRTTRVDFAKNKISIDNFEINNRVSGAGKASSVTLTQASEGITSKNAEISLYD 240  
DB 181 SADRTTRVDFAKNKISIDNFEINNRVSGAGKASSVTLTQASEGITSKNAEISLYD 240

QY 241 GATLNLASSSVKLMGNVWMLQYVAYLAPSYSTINTSKVTGEVNFHNTVGDKNAAQA 300

```
|||||
Db 241 GATNLASSSVKLMGNVWGRLOYGAYLAPSYSTINTSKVTGEVNFHNTVGDKNAAQA 300
QY 301 GIANKKTNIGTLDLWQSLNIIAPPPEGKYDKPNNTPSQSGAKNDKSNKAKNDQESS 360
Db 301 GIANKKTNIGTLDLWQSLNIIAPPPEGKYDKPNNTPSQSGAKNDKSNKAKNDQESS 360
QY 361 QNNSNTQVINPPNSAQTEVQPTQVTDGPFAGGKDTVVNINRINTNADGTIRVGGFKASL 420
Db 361 QNNSNTQVINPPNSAQTEVQPTQVTDGPFAGGKDTVVNINRINTNADGTIRVGGFKASL 420
QY 421 TTNAAHLHIGKGVNLNNOASGRSLIVENLTGNTVDPGLRVNNOVGYALAGSSANFEF 480
Db 421 TTNAAHLHIGKGVNLNNOASGRSLIVENLTGNTVDPGLRVNNOVGYALAGSSANFEF 480
QY 481 KAGTDTKNGTATFNNDISLGRFVNLKVDATNFANFKGIDTNGGFFNTLDFSGYTDKVNINK 540
Db 481 KAGTDTKNGTATFNNDISLGRFVNLKVDATNFANFKGIDTNGGFFNTLDFSGYTDKVNINK 540
QY 541 LITASTNVAVKFNINELIVKTNIGSVGEYTHFSEDIQSQRINTVRLTGTGRSLFSGV 600
Db 541 LITASTNVAVKFNINELIVKTNIGSVGEYTHFSEDIQSQRINTVRLTGTGRSLFSGV 600
QY 601 KFKGGEKLVIDEFYSPWNYFDARNIKNVEITNKLAFPGQSPWGTSKLMFNNTLQONA 660
Db 601 KFKGGEKLVIDEFYSPWNYFDARNIKNVEITNKLAFPGQSPWGTSKLMFNNTLQONA 660
QY 661 VMDYQFSLNLTQGGDFINNOGTINTLVGGKVATLSVGNAAAMFNNDIDSATGYKPLI 720
Db 661 VMDYQFSLNLTQGGDFINNOGTINTLVGGKVATLSVGNAAAMFNNDIDSATGYKPLI 720
QY 721 KINSADLLKTEHVLLKAKIIGYGVSTGTNGISNVNLEQFKERLALYNNNNRMDTCV 780
Db 721 KINSADLLKTEHVLLKAKIIGYGVSTGTNGISNVNLEQFKERLALYNNNNRMDTCV 780
QY 781 VNTDDIKACMAIGQSVNPNPNKYKILIGKAWNIGISKTANGSKISVYVYLGSTPTE 840
Db 781 VNTDDIKACMAIGQSVNPNPNKYKILIGKAWNIGISKTANGSKISVYVYLGSTPTE 840
QY 841 NGGNTNLTPTNTNARSANNAALQAPPAQSPATPNLVAIINHDPCTVIESVFELANRSK 900
Db 841 NGGNTNLTPTNTNARSANNAALQAPPAQSPATPNLVAIINHDPCTVIESVFELANRSK 900
QY 901 DIDTLVANSAGRODLOTLTLDSDHAGYARKMIDATSAEITKOLNTATTTLNNIASLE 960
Db 901 DIDTLVANSAGRODLOTLTLDSDHAGYARKMIDATSAEITKOLNTATTTLNNIASLE 960
QY 961 HTSGLOTLNLSNAMLNLSRLNLSRRHTNHDSPAKRLQALKDQKFALESAAEVLYQF 1020
Db 961 HTSGLOTLNLSNAMLNLSRLNLSRRHTNHDSPAKRLQALKDQKFALESAAEVLYQF 1020
QY 1021 APKYKPTNVWANAIGGTSNANGSNASLYGTSAGVDAYLNGOVEAIVGGFSGYSFSPNN 1080
Db 1021 APKYKPTNVWANAIGGTSNANGSNASLYGTSAGVDAYLNGOVEAIVGGFSGYSFSPNN 1080
QY 1081 RANLSNGANTNFGYGRIFANQHEPDEAQAAGLSDQSSINFKSALLQDLNOSYHYLA 1140
Db 1081 RANLSNGANTNFGYGRIFANQHEPDEAQAAGLSDQSSINFKSALLQDLNOSYHYLA 1140
QY 1141 YSAATRASVGYDFAFRNALVLPKPSGVSYNHLGSTNFKSNSTNQVALKNGSSOHLFNA 1200
Db 1141 YSAATRASVGYDFAFRNALVLPKPSGVSYNHLGSTNFKSNSTNQVALKNGSSOHLFNA 1200
QY 1201 SANVEARYYGGTSTFYNNAGVLOEFAHVGSSNAAASLNTFKVNAARNPLNTHARVMMGGE 1260
Db 1201 SANVEARYYGGTSTFYNNAGVLOEFAHVGSSNAAASLNTFKVNAARNPLNTHARVMMGGE 1260
QY 1261 LKLAKEVLNLGVYVYLNHLNLSNIGHFASNLGMRYSF 1296
Db 1261 LKLAKEVLNLGVYVYLNHLNLSNIGHFASNLGMRYSF 1296
```

RESULT 3

```
US-08-466-662-3
; Sequence 3, Application US/08466662B
; Patent No. 6130059
; GENERAL INFORMATION:
; APPLICANT: Covacci, Antonello
; APPLICANT: Bugnoli, Massimo
; APPLICANT: Telford, John
; APPLICANT: Macchia, Giovanni
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
; TITLE OF INVENTION: Diagnostics
; FILE REFERENCE: CHIR0057
; CURRENT APPLICATION NUMBER: US/08/466,662B
; CURRENT FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1296
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-466-662-3
```

```
Query Match 100.0%; Score 1296; DB 4; Length 1296;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIQOTHRKNRPVLSLALVGALVSIITPOQSHAAFTTVIIPAIVGGIATGTAAGTVSGL 60
Db 1 MEIQOTHRKNRPVLSLALVGALVSIITPOQSHAAFTTVIIPAIVGGIATGTAAGTVSGL 60
QY 61 LSWGLKQAEANKTPDKPKVWRIQAGKGFNEFPKNEYDLYSLSSKIDGGWDGNAAR 120
Db 61 LSWGLKQAEANKTPDKPKVWRIQAGKGFNEFPKNEYDLYSLSSKIDGGWDGNAAR 120
QY 121 HYVKGQGNKLEVDKDAVGTTLISGLRNFPTGGDLVDNMOKATRLRLOFNGNSFTSYKD 180
Db 121 HYVKGQGNKLEVDKDAVGTTLISGLRNFPTGGDLVDNMOKATRLRLOFNGNSFTSYKD 180
QY 181 SADRTTRVDFAKNTSIDNFVEINNRVSGAGRKASSTVLTLOASEGITSIDKNAEISLYD 240
Db 181 SADRTTRVDFAKNTSIDNFVEINNRVSGAGRKASSTVLTLOASEGITSIDKNAEISLYD 240
QY 241 GATNLASSSVKLMGNVWGRLOYGAYLAPSYSTINTSKVTGEVNFHNTVGDKNAAQA 300
Db 241 GATNLASSSVKLMGNVWGRLOYGAYLAPSYSTINTSKVTGEVNFHNTVGDKNAAQA 300
QY 301 GIANKKTNIGTLDLWQSLNIIAPPPEGKYDKPNNTPSQSGAKNDKSNKAKNDQESS 360
Db 301 GIANKKTNIGTLDLWQSLNIIAPPPEGKYDKPNNTPSQSGAKNDKSNKAKNDQESS 360
QY 361 QNNSNTQVINPPNSAQTEVQPTQVTDGPFAGGKDTVVNINRINTNADGTIRVGGFKASL 420
Db 361 QNNSNTQVINPPNSAQTEVQPTQVTDGPFAGGKDTVVNINRINTNADGTIRVGGFKASL 420
QY 421 TTNAAHLHIGKGVNLNNOASGRSLIVENLTGNTVDPGLRVNNOVGYALAGSSANFEF 480
Db 421 TTNAAHLHIGKGVNLNNOASGRSLIVENLTGNTVDPGLRVNNOVGYALAGSSANFEF 480
QY 481 KAGTDTKNGTATFNNDISLGRFVNLKVDATNFANFKGIDTNGGFFNTLDFSGYTDKVNINK 540
Db 481 KAGTDTKNGTATFNNDISLGRFVNLKVDATNFANFKGIDTNGGFFNTLDFSGYTDKVNINK 540
QY 541 LITASTNVAVKFNINELIVKTNIGSVGEYTHFSEDIQSQRINTVRLTGTGRSLFSGV 600
Db 541 LITASTNVAVKFNINELIVKTNIGSVGEYTHFSEDIQSQRINTVRLTGTGRSLFSGV 600
QY 601 KFKGGEKLVIDEFYSPWNYFDARNIKNVEITNKLAFPGQSPWGTSKLMFNNTLQONA 660
Db 601 KFKGGEKLVIDEFYSPWNYFDARNIKNVEITNKLAFPGQSPWGTSKLMFNNTLQONA 660
QY 661 VMDYQFSLNLTQGGDFINNOGTINTLVGGKVATLSVGNAAAMFNNDIDSATGYKPLI 720
|||||
```

Db 661 VMDYSQFSLTIQDFFINQGTINYLVRGKQVATLSVGNAAAMFNNDDISATGFKPLI 720  
QY 721 KNSAODLIKNTHEVLLKAKIIGYGNVSTGTCISNVNLEBOFKERLALYNNNNRMDTCV 780  
Db 721 KNSAODLIKNTHEVLLKAKIIGYGNVSTGTCISNVNLEBOFKERLALYNNNNRMDTCV 780  
QY 781 VRNTDDIKACGMAIGDQSMVNPNDYKYLIGKAWKIGISKTANGSKISVYILGNSTPTE 840  
Db 781 VRNTDDIKACGMAIGDQSMVNPNDYKYLIGKAWKIGISKTANGSKISVYILGNSTPTE 840  
QY 841 NGGNTNLPTNTSNARSANNAQAQAPSAQAPSAQAPSAQAPSAQAPSAQAPSAQAPSA 900  
Db 841 NGGNTNLPTNTSNARSANNAQAQAPSAQAPSAQAPSAQAPSAQAPSAQAPSAQAPSA 900  
QY 901 DIDTLYANSAGORLLQTLIDSHDAGYARKWIDATSAEITKQINTATTTLNNIASLE 960  
Db 901 DIDTLYANSAGORLLQTLIDSHDAGYARKWIDATSAEITKQINTATTTLNNIASLE 960  
QY 961 HKTSGLOTLSLNSAMILNSRLNLSRRHTNHIDSFARKLQALKDOKFASLEAAEVLQF 1020  
Db 961 HKTSGLOTLSLNSAMILNSRLNLSRRHTNHIDSFARKLQALKDOKFASLEAAEVLQF 1020  
QY 1021 APKYKPTNVWANAIGTSLNNGSNASLYGTSGVDAYLINGQVEALVGGFGSYGSFNN 1080  
Db 1021 APKYKPTNVWANAIGTSLNNGSNASLYGTSGVDAYLINGQVEALVGGFGSYGSFNN 1080  
QY 1081 RANSLNGANTNFGVSRIFANOHEFDEAOCALGSDQSSLNFKSALLQDLNQSYHILA 1140  
Db 1081 RANSLNGANTNFGVSRIFANOHEFDEAOCALGSDQSSLNFKSALLQDLNQSYHILA 1140  
QY 1141 YSAATRASVGYDFAFRNALVLPKPSGVSYNHLGSTNFKSTNOVALKNGSSOHLFNA 1200  
Db 1141 YSAATRASVGYDFAFRNALVLPKPSGVSYNHLGSTNFKSTNOVALKNGSSOHLFNA 1200  
QY 1201 SANVEARYYGTSTFYNNAGVQLOFAHVGSNNAASLNTFKYNAARNPLNTHARYMGMGE 1260  
Db 1201 SANVEARYYGTSTFYNNAGVQLOFAHVGSNNAASLNTFKYNAARNPLNTHARYMGMGE 1260  
QY 1261 LKLAKEVFLNLGVVYLNHLNLSNIGHFASNLGMRYSF 1296  
Db 1261 LKLAKEVFLNLGVVYLNHLNLSNIGHFASNLGMRYSF 1296

RESULT 4  
US-08-200-232-2  
; Sequence 2, Application US/08200232  
; Patent No. 5721349  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI  
; TITLE OF INVENTION: AND RELATED METHODS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: NEEDLE & ROSENBERG P.C.  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/200,232  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 2200.023

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1287 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-200-232-2

Query Match 7.6%; Score 99; DB 1; Length 1287;  
Best Local Similarity 100.0%; Pred. No. 8.3e-89;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 697 VGNAAAMFNNDDISATGFKPLIKINSADLIKNTHEVLLKAKIIGYGNVSTGTCISN 756  
Db 689 VGNAAAMFNNDDISATGFKPLIKINSADLIKNTHEVLLKAKIIGYGNVSTGTCISN 748  
QY 757 VNLEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 795  
Db 749 VNLEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 787

RESULT 5  
PCT-US95-02219-2  
; Sequence 2, Application PC/TUS9502219  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI  
; TITLE OF INVENTION: AND RELATED METHODS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG P.C.  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02219  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 2200.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1287 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-02219-2

Query Match 7.6%; Score 99; DB 5; Length 1287;  
Best Local Similarity 100.0%; Pred. No. 8.3e-89;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 697 VGNAAAMFNNDDISATGFKPLIKINSADLIKNTHEVLLKAKIIGYGNVSTGTCISN 756  
Db 689 VGNAAAMFNNDDISATGFKPLIKINSADLIKNTHEVLLKAKIIGYGNVSTGTCISN 748

OY 757 VNLEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 795  
Db 749 VNLEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 787

RESULT 6  
PCT-US95-02219A-2  
; Sequence 2, Application PC/TUS9502219A  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Tummuru, Murali KR  
; APPLICANT: Cao, Ping  
; APPLICANT: Thompson, Stuart A.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02219A  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 2200.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1287 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-02219A-2

Query Match 7.6%; Score 99; DB 5; Length 1287;  
Best Local Similarity 100.0%; Pred. No. 8.3e-89;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 697 VGNAAAMFNDDISATGYFKPLKINSADLIKNTHVLKAKIIGYGVSTGTNGISN 756  
Db 689 VGNAAAMFNDDISATGYFKPLKINSADLIKNTHVLKAKIIGYGVSTGTNGISN 748  
OY 757 VNLEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 795  
Db 749 VNLEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 787

RESULT 7  
US-08-295-643-2  
; Sequence 2, Application US/08295643  
; Patent No. 5859219  
; GENERAL INFORMATION:  
; APPLICANT: COVER, TIMOTHY L.  
; APPLICANT: BLASER, MARTIN J.  
; TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
; HELICOBACTER PYLORI AND METHODS TO USE SAME  
; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
; STREET: Suite 1200, 127 Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/295,643  
; FILING DATE: 26-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SPRATT, GWENDOLYN D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 2200.025  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-295-643-2

Query Match 1.8%; Score 23; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-15;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 34 AFTTIVIIIPAIVGGIATGTAVGT 56  
Db 1 AFTTIVIIIPAIVGGIATGTAVGT 23

RESULT 8  
US-08-473-265-1  
; Sequence 1, Application US/08473265  
; Patent No. 6013463  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
; HELICOBACTER PYLORI AND METHODS TO USE THE SAME  
; FILE REFERENCE: 22000.0026  
; CURRENT APPLICATION NUMBER: US/08/473,265  
; CURRENT FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/284,747  
; EARLIER FILING DATE: 1994-08-02  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence./No. 6013463e -  
; OTHER INFORMATION: synthetic construct  
; US-08-473-265-1

Query Match 1.8%; Score 23; DB 3; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-15;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 34 AFTTIVIIIPAIVGGIATGTAVGT 56

Db 1 AFETTIIIPAIVGGIATGTAAGT 23  
|||||

## RESULT 9

US-08-284-747-1  
; Sequence 1, Application US/08284747A  
; Patent No. 6054132  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
; FILE OF INVENTION: HELICOBLASTER PYLORI AND METHODS TO USE THE SAME  
; FILE REFERENCE: 22000.0026  
; CURRENT APPLICATION NUMBER: US/08/284,747A  
; CURRENT FILING DATE: 1994-08-02  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: No. 6054132e -  
US-08-284-747-1

Query Match 1.8%; Score 23; DB 3; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-15;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AFETTIIIPAIVGGIATGTAAGT 56  
|||||

Db 1 AFETTIIIPAIVGGIATGTAAGT 23  
|||||

## RESULT 10

US-08-200-232-4  
; Sequence 4, Application US/08200232  
; Patent No. 5721349  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI  
; TITLE OF INVENTION: AND RELATED METHODS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG P.C.  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Spratt, Gwendolyn D.  
REGISTRATION NUMBER: 36,016  
REFERENCE/DOCKET NUMBER: 2200.023  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 513 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-200-232-4

Query Match 1.5%; Score 20; DB 1; Length 513;  
Best Local Similarity 100.0%; Pred. No. 3.5e-11;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
|||||

Db 44 OPTQVIDGPFAGGKDTVVNI 63  
|||||

## RESULT 11

PCT-US95-02219-4  
; Sequence 4, Application PC/TUS9502219  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI  
; TITLE OF INVENTION: AND RELATED METHODS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG P.C.  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Spratt, Gwendolyn D.  
REGISTRATION NUMBER: 36,016  
REFERENCE/DOCKET NUMBER: 2200.023  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 513 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-02219-4

Query Match 1.5%; Score 20; DB 5; Length 513;  
Best Local Similarity 100.0%; Pred. No. 3.5e-11;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
|||||

Db 44 OPTQVIDGPFAGGKDTVVNI 63  
|||||

## RESULT 12

PCT-US95-02219A-4  
; Sequence 4, Application PC/TUS9502219A  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Tummuru, Murali KR  
; APPLICANT: Cao, Ping  
; APPLICANT: Thompson, Stuart A.

APPLICANT: Blaser, Martin J.  
TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEEDLE & ROSENBERG P.C.  
STREET: 127 Peachtree Street, Suite 1200  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: COVER, TIMOTHY L.  
TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
NUMBER OF SEQUENCES: 22  
FILING DATE: 26-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: SPRATT, GWENDOLYN D.  
REGISTRATION NUMBER: 36,016  
REFERENCE/DOCKET NUMBER: 2200.023  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 513 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-02219A-4

Query Match 1.5%; Score 20; DB 5; Length 513;  
Best Local Similarity 100.0%; Pred. No. 3.5e-11;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 QPTQVIDGPFAGGKDTVVNI 400  
|||||  
DB 44 QPTQVIDGPFAGGKDTVVNI 63

RESULT 13  
US-08-295-643-15  
Sequence 15, Application US/08295643  
Patent No. 5859219  
GENERAL INFORMATION:  
APPLICANT: COVER, TIMOTHY L.  
ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
STREET: Suite 1200, 127 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: COVER, TIMOTHY L.  
TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
NUMBER OF SEQUENCES: 22  
FILING DATE: 26-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: SPRATT, GWENDOLYN D.

REGISTRATION NUMBER: 36,016  
REFERENCE/DOCKET NUMBER: 2200.025  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-295-643-15

Query Match 1.1%; Score 14; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 LGQFNGNSFTSYKD 180  
|||||  
DB 1 LGQFNGNSFTSYKD 14

RESULT 14  
US-08-295-643-17  
Sequence 17, Application US/08295643  
Patent No. 5859219  
GENERAL INFORMATION:  
APPLICANT: COVER, TIMOTHY L.  
ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
STREET: Suite 1200, 127 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: COVER, TIMOTHY L.  
TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
NUMBER OF SEQUENCES: 22  
FILING DATE: 26-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: SPRATT, GWENDOLYN D.  
REGISTRATION NUMBER: 36,016  
REFERENCE/DOCKET NUMBER: 2200.025  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 3  
OTHER INFORMATION: /note="EITHER VALINE OR ISOLEUCINE CAN BE USED  
HERE"  
US-08-295-643-17

Query Match 0.8%; Score 10; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 TRVDNNAKNI 195  
Db 1 TRVDNNAKNI 10

RESULT 15  
US-08-295-643-16  
; Sequence 16, Application US/08295643  
; Patent No. 5859219  
; GENERAL INFORMATION:  
; APPLICANT: COVER, TIMOTHY L.  
; APPLICANT: BLASER, MARTIN J.  
; TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
; TITLE OF INVENTION: HELICOBACTER PYLORI AND METHODS TO USE SAME  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
; STREET: Suite 1200, 127 Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/295,643  
; APPLICATION DATE: 26-AUG-1994  
; FILING DATE: 26-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SPRATT, GWENDOLYN D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 2200.025  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-295-643-16

Query Match 0.6%; Score 8; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 625 NIKNVEIT 632  
Db 1 NIKNVEIT 8

RESULT 16  
PCT-US95-04121-43  
; Sequence 43, Application PC/TUS9504121  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Haptenated Peptides and Uses Thereof  
; NUMBER OF SEQUENCES: 62  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04121  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/222,206  
FILING DATE: April 1, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Vanstone, Darlene A.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 079.2PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6010  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
PCT-US95-04121-43

Query Match 0.5%; Score 7; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQCAL 1115  
Db 1 FEAQCAL 7

RESULT 17  
PCT-US95-04121-46  
; Sequence 46, Application PC/TUS9504121  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Haptenated Peptides and Uses Thereof  
; NUMBER OF SEQUENCES: 62  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04121  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/222,206  
FILING DATE: April 1, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Vanstone, Darlene A.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 079.2PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6010  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
PCT-US95-04121-46

Query Match 0.5%; Score 7; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1109 FEAQAL 1115  
1111111  
Db 1 FEAQAL 7

## RESULT 18

US-08-337-646A-24  
; Sequence 24, Application US/08337646A  
; Patent No. 5856171  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/337,646A  
; FILING DATE: 10-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/248,819  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/112,208  
; FILING DATE: 26-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000620  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-337-646A-24

Query Match 0.5%; Score 7; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 SLALVGA 22  
1111111  
Db 2 SLALVGA 8

## RESULT 19

US-08-927-326-24  
; Sequence 24, Application US/08927326  
; Patent No. 6184202  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue

; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/927,326  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/337,646  
; FILING DATE: 10-NOV-1994  
; APPLICATION NUMBER: US 08/248,819  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/112,208  
; FILING DATE: 26-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000620  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-927-326-24

Query Match 0.5%; Score 7; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 SLALVGA 22  
1111111  
Db 2 SLALVGA 8

## RESULT 20

US-08-207-481-5  
; Sequence 5, Application US/08207481  
; Patent No. 5820866  
; GENERAL INFORMATION:  
; APPLICANT: Kappler, John W.  
; APPLICANT: Maritack, Philippa  
; TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL  
; TITLE OF INVENTION: REGULATION  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
; STREET: 1700 LINCOLN STREET, SUITE 3500  
; CITY: DENVER  
; STATE: COLORADO  
; COUNTRY: USA  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/207,481  
; FILING DATE: 04-MAR-1994  
; CLASSIFICATION: 435



ATTORNEY/AGENT INFORMATION:  
NAME: Kovarik, Joseph E.  
REGISTRATION NUMBER: 33,005  
REFERENCE/DOCKET NUMBER: 2879-8  
TELEPHONE: 303/863-9700  
TELEFAX: 303/863-0223  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-207-481-5

Query Match 0.5%; Score 7; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
Db 3 FEAQAL 9

RESULT 21  
US-08-640-344-10  
Sequence 10, Application US/08640344  
Patent No. 5824315  
GENERAL INFORMATION:  
APPLICANT: NAG, BISHWAJIT  
APPLICANT: MUKKU, PRABHA  
APPLICANT: DESHPANDE, SHRIKANT  
TITLE OF INVENTION: IMPROVING BINDING AFFINITY OF ANTIGENIC  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP  
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/640,344  
FILING DATE: 30-APR-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: STORELLA ESQ., JOHN R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 14058-004800  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-640-344-10

Query Match 0.5%; Score 7; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
Db 3 FEAQAL 9

RESULT 22  
US-08-640-344-11  
Sequence 11, Application US/08640344  
Patent No. 5824315  
GENERAL INFORMATION:  
APPLICANT: NAG, BISHWAJIT  
APPLICANT: MUKKU, PRABHA  
APPLICANT: DESHPANDE, SHRIKANT  
TITLE OF INVENTION: IMPROVING BINDING AFFINITY OF ANTIGENIC  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP  
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/640,344  
FILING DATE: 30-APR-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: STORELLA ESQ., JOHN R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 14058-004800  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-640-344-11

Query Match 0.5%; Score 7; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
Db 3 FEAQAL 9

RESULT 23  
PCT-US95-02689-5  
Sequence 5, Application PC/TUS9502689  
GENERAL INFORMATION:  
APPLICANT: National Jewish Center for Immunology and  
APPLICANT: Respiratory Medicine  
APPLICANT: Kappler, John W.  
APPLICANT: Marrack, Philippa  
TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL REGULATION  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
STREET: 1700 LINCOLN STREET, SUITE 3500  
CITY: DENVER

STATE: COLORADO  
COUNTRY: USA  
ZIP: 80202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02689  
FILING DATE: 03-MAR-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kovarik, Joseph E.  
REGISTRATION NUMBER: 33,005  
REFERENCE/DOCKET NUMBER: 2879-8-PCT  
TELEPHONE: 303/863-9700  
TELEFAX: 303/863-0223  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-02689-5

Query Match 0.5%; Score 7; DB 5; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 FEAQAL 1115  
Db 3 FEAQAL 9

RESULT 24  
US-08-418-893D-8  
Sequence 8, Application US/08418893D  
Patent No. 5559220  
GENERAL INFORMATION:  
APPLICANT: ROESSLER, PAUL G  
APPLICANT: OHLROGGE, JOHN B  
TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A  
TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTILLA CRYPTICA  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY  
STREET: 1617 Cole Blvd.  
CITY: Golden  
STATE: CO  
COUNTRY: USA  
ZIP: 80401-3393  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/418,893D  
FILING DATE: April 7, 1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/104,938  
FILING DATE: September 14, 1993  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: O'CONNOR, EDNA  
REGISTRATION NUMBER: 29,252  
REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 303-231-1000  
TELEFAX: 303-231-1098  
TELEX:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
US-08-418-893D-8

Query Match 0.5%; Score 7; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 AIVGGIA 49  
Db 3 AIVGGIA 9

RESULT 25  
US-09-007-905-59  
Sequence 59, Application US/09007905  
Patent No. 6221585  
GENERAL INFORMATION:  
APPLICANT: Fran ois J.-M. Iris and Jean-Louis Pourny  
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES UNDERLYING  
TITLE OF INVENTION: DEFINED PHENOTYPES  
NUMBER OF SEQUENCES: 70  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/007,905  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Antler, Adriane M.  
REGISTRATION NUMBER: 32,605  
REFERENCE/DOCKET NUMBER: 9408-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-007-905-59

Query Match 0.5%; Score 7; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1165 SVGVSYN 1171

Db 4 SVGVSYN 10  
|||||

RESULT 26  
US-09-007-905-64  
; Sequence 64, Application US/09007905  
; Patent No. 6221585  
; GENERAL INFORMATION:  
; APPLICANT: Fran ois J.-M. Iris and Jean-Louis Pourny  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES UNDERLYING  
; TITLE OF INVENTION: DEFINED PHENOTYPES  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/007,905  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Antler, Adriane M.  
; REGISTRATION NUMBER: 32,605  
; REFERENCE/DOCKET NUMBER: 9408-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-09-007-905-64

Query Match 0.5%; Score 7; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1165 SVGVSYN 1171  
Db 4 SVGVSYN 10  
|||||

RESULT 27  
US-09-166-028-7  
; Sequence 7, Application US/09166028  
; Patent No. 6245885  
; GENERAL INFORMATION:  
; APPLICANT: Gordon C. Shore et al.  
; TITLE OF INVENTION: BAX-MEDIATED APOPTOSIS MODULATING  
; TITLE OF INVENTION: REAGENTS AND METHODS  
; FILE REFERENCE: 50013/011001  
; CURRENT APPLICATION NUMBER: US/09/166.028  
; CURRENT FILING DATE: 1998-10-05  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-166-028-7

Query Match 0.5%; Score 7; DB 4; Length 22;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 16 SLALVGA 22  
Db 5 SLALVGA 11  
|||||

RESULT 28  
US-08-428-091-1  
; Sequence 1, Application US/08428091  
; Patent No. 5683890  
; GENERAL INFORMATION:  
; APPLICANT: GERMOND, JACQUES.-EDOUARD  
; APPLICANT: MARCISET, OLIVIER  
; APPLICANT: MOLLET, BEAT  
; TITLE OF INVENTION: BACTERIOICINS OF STREPTOCOCCUS  
; TITLE OF INVENTION: THERMOPHILUS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS  
; STREET: 1155 AVENUE OF THE AMERICA  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,091  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP PCT/EP94/02805  
; FILING DATE: 24-AUG-1994  
; APPLICATION NUMBER: CH 2628/93-7  
; FILING DATE: 03-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FANUCCI, ALLAN A  
; REGISTRATION NUMBER: 30256  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790 9090  
; TELEFAX: 212 869 8864  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 62 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus thermophilus  
; STRAIN: CNCM I-1351  
US-08-428-091-1

Query Match 0.5%; Score 7; DB 1; Length 62;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 467 GGYALAG 473  
Db 11 GGYALAG 17  
|||||

RESULT 29  
US-09-007-905-36

; Sequence 36, Application US/09007905  
; Patent No. 6221585  
; GENERAL INFORMATION:  
; APPLICANT: Fran Ols J.-M. Iris and Jean-Louis Pourny  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES UNDERLYING  
; TITLE OF INVENTION: DEFINED PHENOTYPES  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/007,905  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Antler, Adriane M.  
; REGISTRATION NUMBER: 32,605  
; REFERENCE/DOCKET NUMBER: 9408-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-869-9090  
; TELEX: 66141 PENNIE  
; TELEFAX: 212-869-8864  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 65 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-09-007-905-36

Query Match 0.5%; Score 7; DB 4; Length 65;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1165 SVGVSYN 1171  
Db 17 SVGVSYN 23

RESULT 30  
US-09-007-905-40  
; Sequence 40, Application US/09007905  
; Patent No. 6221585  
; GENERAL INFORMATION:  
; APPLICANT: Fran Ols J.-M. Iris and Jean-Louis Pourny  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES UNDERLYING  
; TITLE OF INVENTION: DEFINED PHENOTYPES  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/007,905

; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Antler, Adriane M.  
; REGISTRATION NUMBER: 32,605  
; REFERENCE/DOCKET NUMBER: 9408-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 66 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-09-007-905-40

Query Match 0.5%; Score 7; DB 4; Length 66;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1165 SVGVSYN 1171  
Db 18 SVGVSYN 24

RESULT 31  
US-08-207-481-45  
; Sequence 45, Application US/08207481  
; Patent No. 5820866  
; GENERAL INFORMATION:  
; APPLICANT: Kappler, John W.  
; APPLICANT: Matrack, Philippa  
; TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL  
; TITLE OF INVENTION: REGULATION  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
; STREET: 1700 LINCOLN STREET, SUITE 3500  
; CITY: DENVER  
; STATE: COLORADO  
; COUNTRY: USA  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/207,481  
; FILING DATE: 04-MAR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kovarik, Joseph E.  
; REGISTRATION NUMBER: 33,005  
; REFERENCE/DOCKET NUMBER: 2879-8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 67 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-207-481-45

Query Match 0.5%; Score 7; DB 2; Length 67;  
Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 FEAQAL 1115  
Db 34 FEAQAL 40

## RESULT 32

PCT-US95-02689-52  
; Sequence 52, Application PC/TUS9502689  
; GENERAL INFORMATION:  
; APPLICANT: National Jewish Center for Immunology and  
; APPLICANT: Respiratory Medicine  
; APPLICANT: Kappler, John W.  
; APPLICANT: Marrack, Philippa  
; TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL REGULATION  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
; STREET: 1700 LINCOLN STREET, SUITE 3500  
; CITY: DENVER  
; STATE: COLORADO  
; COUNTRY: USA  
; ZIP: 80202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

APPLICATION NUMBER: PCT/US95/02689  
FILING DATE: 03-MAR-1995

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
; NAME: Kovarik, Joseph E.  
; REGISTRATION NUMBER: 33,005  
; REFERENCE/DOCKET NUMBER: 2879-8-PCT  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223

INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
; LENGTH: 67 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

MOLECULE TYPE: protein  
PCT-US95-02689-52

Query Match 0.5%; Score 7; DB 5; Length 67;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 FEAQAL 1115  
Db 34 FEAQAL 40

## RESULT 33

US-09-007-905-30  
; Sequence 30, Application US/09007905  
; Patent No. 6221585  
; GENERAL INFORMATION:  
; APPLICANT: Fran Ols J.-M. Iris and Jean-Louis Pourny  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES UNDERLYING  
; TITLE OF INVENTION: DEFINED PHENOTYPES  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA

ZIP: 10036-2711  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 2.0  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/007,905  
; FILING DATE:

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
; NAME: Antler, Adriane M.  
; REGISTRATION NUMBER: 32,605  
; REFERENCE/DOCKET NUMBER: 9408-003  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 69 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-09-007-905-30

Query Match 0.5%; Score 7; DB 4; Length 69;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1165 SVGVSYN 1171  
Db 22 SVGVSYN 28

## RESULT 34

US-08-428-091-4  
; Sequence 4, Application US/08428091  
; Patent No. 5683890  
; GENERAL INFORMATION:  
; APPLICANT: GERMOND, JACQUES.-EDOUARD  
; APPLICANT: MOLLET, OLIVIER  
; APPLICANT: MOLLET, BEAT  
; TITLE OF INVENTION: BACTERIOGINS OF STREPTOCOCCUS  
; TITLE OF INVENTION: THERMOPHILUS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS  
; STREET: 1155 AVENUE OF THE AMERICA  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,091  
; FILING DATE:  
; CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP PCT/EP94/02805  
; FILING DATE: 24-AUG-1994  
; APPLICATION NUMBER: CH 2628/93-7  
; FILING DATE: 03-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FANUCCI, ALLAN A  
; REGISTRATION NUMBER: 30256  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790 9090  
 TELEFAX: 212 869 8864  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 85 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-428-091-4

Query Match 0.5%; Score 7; DB 1; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 467 GGYALAG 473  
 |||||  
 DB 34 GGYALAG 40

RESULT 35  
 5514582-19  
 Patent No. 5514582  
 APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
 TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
 IMMUNOGLOBULINS  
 NUMBER OF SEQUENCES: 43  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/185,670  
 FILING DATE: 21-JAN-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 986,931  
 FILING DATE: 08-DEC-1992  
 APPLICATION NUMBER: 808,122  
 FILING DATE: 16-DEC-1991  
 APPLICATION NUMBER: 440,625  
 FILING DATE: 22-NOV-1989  
 APPLICATION NUMBER: 315,015  
 FILING DATE: 23-FEB-1989  
 SEQ ID NO: 19:  
 LENGTH: 123  
 5514582-19

Query Match 0.5%; Score 7; DB 6; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 NNPDNYK 807  
 |||||  
 DB 87 NNPDNYK 93

RESULT 36  
 US-08-077-848A-3  
 Sequence 3, Application US/08077848A  
 Patent No. 5470955  
 GENERAL INFORMATION:  
 APPLICANT: Craig, Ruth W.  
 TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Spensley Horn Jubas & Lubitz  
 STREET: 1880 Century Park East, Suite 500  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: USA  
 ZIP: 90067  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/077,848A  
 FILING DATE: 16-JUN-1993  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile, Ph.D., Lisa A.  
 REGISTRATION NUMBER: 38,347  
 REFERENCE/DOCKET NUMBER: PD-2845  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 455-5100  
 TELEFAX: (619) 455-5110  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 154 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 IMMEDIATE SOURCE:  
 CLONE: bcl-2alpha  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..154  
 US-08-077-848A-3

Query Match 0.5%; Score 7; DB 1; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 |||||  
 DB 137 SLALVGA 143

RESULT 37  
 US-09-211-640-3  
 Sequence 3, Application US/09211640  
 Patent No. 6020466  
 GENERAL INFORMATION:  
 APPLICANT: Craig, Ruth W.  
 TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Spensley Horn Jubas & Lubitz  
 STREET: 1880 Century Park East, Suite 500  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: USA  
 ZIP: 90067  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/211,640  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/441,375  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile, Ph.D., Lisa A.  
 REGISTRATION NUMBER: 38,347  
 REFERENCE/DOCKET NUMBER: PD-2845  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 455-5100  
 TELEFAX: (619) 455-5110  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 154 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
IMMEDIATE SOURCE:  
CLONE: bcl-2alpha  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..154  
US-09-211-640-3

Query Match 0.5%; Score 7; DB 3; Length 154;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 137 SLALVGA 143

RESULT 38  
US-09-378-536-3  
; Sequence 3, Application US/09378536  
; Patent No. 6200763  
; GENERAL INFORMATION:  
; APPLICANT: Craig, Ruth W.  
; TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1  
; MEDIUM TYPE: POLYPEPTIDE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Juba & Lubitz  
; STREET: 1880 Century Park East, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90067  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/378,536  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/077,848  
; FILING DATE: 16-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: PD-2845  
; TELEPHONE: (619) 455-5100  
; TELEFAX: (619) 455-5110  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 154 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; IMMEDIATE SOURCE:  
; CLONE: bcl-2alpha  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..154  
US-09-378-536-3

Query Match 0.5%; Score 7; DB 4; Length 154;

Best Local Similarity 100.0%; Pred. No. 76;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 137 SLALVGA 143

RESULT 39  
US-08-960-190A-38  
; Sequence 38, Application US/08960190A  
; Patent No. 6232445  
; GENERAL INFORMATION:  
; APPLICANT: Rhode, Peter R.  
; APPLICANT: Acevedo, Jorge  
; APPLICANT: Burkhardt, Martin  
; APPLICANT: Jiao, Jin-an  
; APPLICANT: Wong, Jing C.  
; TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND  
; METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: usa  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/960,190A  
; FILING DATE: 29-OCT-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Corless, Peter F.  
; REGISTRATION NUMBER: 33,860  
; REFERENCE/DOCKET NUMBER: 48002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-960-190A-38

Query Match 0.5%; Score 7; DB 4; Length 192;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQGAL 1115  
Db 54 FEAQAL 60

RESULT 40  
US-08-557-146-13  
; Sequence 13, Application US/08557146  
; Patent No. 5834290  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn

APPLICANT: Hansson, Lennart  
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic Enzyme (SCCE)  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case, Patent Department  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sturner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
US-08-557-146-13

Query Match 0.5%; Score 7; DB 2; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 ANKTPDK 77  
| | | | | | | |  
DB 133 ANKTPDK 139

RESULT 41  
US-08-408-095-19  
Sequence 19, Application US/08408095  
Patent No. 5858678  
GENERAL INFORMATION:  
APPLICANT: Chinadurai, Govindaswamy  
TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MILO, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/408,095  
FILING DATE: 21-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mack, Susan J.  
REGISTRATION NUMBER: 30,951

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-408-095-19

Query Match 0.5%; Score 7; DB 2; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
| | | | | | | |  
DB 212 SLALVGA 218

RESULT 42  
US-09-154-344-13  
Sequence 13, Application US/09154344  
Patent No. 5981256  
GENERAL INFORMATION:  
APPLICANT: Egelrud, Torbjorn  
APPLICANT: Hansson, Lennart  
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic Enzyme (SCCE)  
TITLE OF INVENTION: Enzyme (SCCE)  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case, Patent Department  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,344  
FILING DATE: 16-SEP-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sturner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
US-09-154-344-13

Query Match 0.5%; Score 7; DB 2; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 71 ANKTPDK 77  
Db 133 ANKTPDK 139

## RESULT 43

US-08-470-535-12  
; Sequence 12, Application US/08470535  
; Patent No. 6050587  
; GENERAL INFORMATION:  
; APPLICANT: Rhodes, Eric T  
; APPLICANT: Nag, Bishwajit  
; TITLE OF INVENTION: PROKARYOTIC EXPRESSION OF MHC PROTEINS  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: One Market Plaza, Steuart Street Tower  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105-1492  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,535  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/143,575  
; FILING DATE: 25-OCT-1993  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-470-535-12

Query Match 0.5%; Score 7; DB 3; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
Db 54 FEAQAL 60

## RESULT 44

US-08-944-483-62  
; Sequence 62, Application US/08944483  
; Patent No. 6232456  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STEWART, KENT D.

; APPLICANT: STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/944,483  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Becker, Cheryl L.  
; REGISTRATION NUMBER: 35,441  
; REFERENCE/DOCKET NUMBER: 6183.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847/935-1729  
; TELEFAX: 847/938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 62:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 230 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6232456e  
US-08-944-483-62

Query Match 0.5%; Score 7; DB 4; Length 230;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 ANKTPDK 77  
Db 133 ANKTPDK 139

## RESULT 45

US-09-027-337-6  
; Sequence 6, Application US/09027337B  
; Patent No. 5972616  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotooshi  
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease Overexpressed in  
; TITLE OF INVENTION: Breast and Ovarian Carcinomas  
; FILE REFERENCE: D6064  
; CURRENT APPLICATION NUMBER: US/09/027,337B  
; CURRENT FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 13  
; SEQ ID NO 6  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of chymotrypsin (Chymb)  
; OTHER INFORMATION: homologous to similar domain in TAGD-15  
US-09-027-337-6

Query Match 0.5%; Score 7; DB 2; Length 231;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 ANKTPDK 77  
Db 134 ANKTPDK 140

## RESULT 46

US-08-408-095-17  
; Sequence 17, Application US/08408095  
; Patent No. 5858678  
; GENERAL INFORMATION:  
; APPLICANT: Chinnadurai, Govindaswamy  
; TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20037

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/408,095  
; FILING DATE: 21-MAR-1995  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: Mack, Susan J.  
; REGISTRATION NUMBER: 30,951  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-7860  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 232 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-408-095-17

Query Match 0.5%; Score 7; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 215 SLALVGA 221

## RESULT 47

US-08-408-095-18  
; Sequence 18, Application US/08408095  
; Patent No. 5858678  
; GENERAL INFORMATION:  
; APPLICANT: Chinnadurai, Govindaswamy  
; TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA

ZIP: 20037  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/408,095  
; FILING DATE: 21-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mack, Susan J.  
; REGISTRATION NUMBER: 30,951  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-7860  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 232 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-408-095-18

Query Match 0.5%; Score 7; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 215 SLALVGA 221

## RESULT 48

US-08-607-269-21  
; Sequence 21, Application US/08607269  
; Patent No. 5702897  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Sato, Takaaki  
; TITLE OF INVENTION: Interaction of Proteins Involved in a  
; TITLE OF INVENTION: Cell Death Pathway  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/607,269  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/226,876  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 9882  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 236 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-607-269-21

Query Match 0.5%; Score 7; DB 1; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
| | | | | | |  
Db 219 SLALVGA 225

## RESULT 49

US-08-607-269-22  
; Sequence 22, Application US/08607269  
; Patent No. 5702897  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Sato, Takaaki  
; TITLE OF INVENTION: Interaction of Proteins Involved in a  
; TITLE OF INVENTION: Cell Death Pathway  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/607,269  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/226,876  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 9882  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-607-269-22

Query Match 0.5%; Score 7; DB 1; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
| | | | | | |  
Db 219 SLALVGA 225

## RESULT 50

US-09-127-048-9  
; Sequence 9, Application US/09127048  
; Patent No. 6165732  
; GENERAL INFORMATION:

; APPLICANT: Korsmeyer, Stanley J.  
; APPLICANT: Schlesinger, Paul H.  
; TITLE OF INVENTION: Method for Identifying Apoptosis Modulating Compounds  
; FILE REFERENCE: 6029-6052  
; CURRENT APPLICATION NUMBER: US/09/127,048  
; CURRENT FILING DATE: 1998-07-31  
; EARLIER APPLICATION NUMBER: 60/061,823  
; EARLIER FILING DATE: 1997-10-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-127-048-9

Query Match 0.5%; Score 7; DB 4; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
| | | | | | |  
Db 219 SLALVGA 225

## RESULT 51

PCT-US95-04600-21  
; Sequence 21, Application PC/TUS9504600  
; GENERAL INFORMATION:  
; APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION  
; TITLE OF INVENTION: Interaction of Proteins Involved in  
; TITLE OF INVENTION: a Cell Death Pathway  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04600  
; FILING DATE: 12-APR-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Imbra, Richard J.  
; REGISTRATION NUMBER: 37,643  
; REFERENCE/DOCKET NUMBER: PP-LJ 1361  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
PCT-US95-04600-21

Query Match 0.5%; Score 7; DB 5; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
| | | | | | |  
Db 219 SLALVGA 225

RESULT 52  
PCT-US95-04600-22  
; Sequence 22, Application PC/TUS9504600  
; GENERAL INFORMATION:  
; APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION  
; TITLE OF INVENTION: Interaction of Proteins Involved in  
; TITLE OF INVENTION: a Cell Death Pathway  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04600  
; FILING DATE: 12-APR-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Imbra, Richard J.  
; REGISTRATION NUMBER: 37,643  
; REFERENCE/DOCKET NUMBER: FP-LJ 1361  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; PCT-US95-04600-22

Query Match 0.5%; Score 7; DB 5; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 16 SLALVGA 22  
|111111|  
Db 219 SLALVGA 225

RESULT 53  
US-08-333-565-51  
; Sequence 51, Application US/08333565  
; Patent No. 5622852  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: Bcl-x/Bcl-2 ASSOCIATED CELL DEATH  
; TITLE OF INVENTION: REGULATOR  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333,565  
; FILING DATE: 31-OCT-1994

; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000700  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-333-565-51

Query Match 0.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 16 SLALVGA 22  
|111111|  
Db 222 SLALVGA 228

RESULT 54  
US-08-112-208C-10  
; Sequence 10, Application US/08112208C  
; Patent No. 5691179  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/112,208C  
; FILING DATE: 26-AUG-1993  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000610  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-112-208C-10

Query Match 0.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 16 SLALVGA 22

Db 222 SLALVGA 228  
|||||

## RESULT 55

US-08-248-819A-10  
; Sequence 10, Application US/08248819A  
; Patent No. 5700638  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/248,819A  
; FILING DATE: 25-NAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/112,208  
; FILING DATE: 26-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000610  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-248-819A-10

Query Match 0.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22  
|||||  
Db 222 SLALVGA 228

## RESULT 56

US-08-248-819A-12  
; Sequence 12, Application US/08248819A  
; Patent No. 5700638  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/248,819A  
; FILING DATE: 25-NAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/112,208  
; FILING DATE: 26-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000610  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-248-819A-12

Query Match 0.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22  
|||||  
Db 222 SLALVGA 228

## RESULT 57

US-08-607-269-20  
; Sequence 20, Application US/08607269  
; Patent No. 5702897  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Sato, Takaaki  
; TITLE OF INVENTION: Interaction of Proteins Involved in a  
; TITLE OF INVENTION: Cell Death Pathway  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/607,269  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/226,876  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 9882  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-607-269-20

Query Match 0.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22  
Db 222 SLALVGA 228

## RESULT 58

US-08-471-058-12  
Sequence 12, Application US/08471058  
Patent No. 5770443  
GENERAL INFORMATION:  
APPLICANT: Kiefer, Michael C.  
APPLICANT: Barr, Philip J.  
TITLE OF INVENTION: NOVEL APOPTOSIS MODULATING  
TITLE OF INVENTION: PROTEINS, DNA ENCODING THE PROTEINS AND METHODS OF USE  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,058  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/320,157

FILING DATE: 07-OCT-1994  
APPLICATION NUMBER: 08/160,067  
FILING DATE: 30-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lehnhardt, Susan K.  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 23647-20007.12  
TELEPHONE: 415-813-5600  
TELEFAX: 415-494-0792  
TELEX: 706141

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-471-058-12

Query Match 0.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22  
Db 222 SLALVGA 228

## RESULT 59

US-08-405-702A-12  
Sequence 12, Application US/08405702A  
Patent No. 5789389  
GENERAL INFORMATION:  
APPLICANT: Tarasewicz, Dariusz G  
APPLICANT: Schott, Brigitte  
APPLICANT: Holzmayer, Tatiana A.  
APPLICANT: Roninson, Igor B.  
TITLE OF INVENTION: BCL2 Derived Genetic Elements Associated  
TITLE OF INVENTION: with Sensitivity to Chemotherapeutic Drugs  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,702A  
FILING DATE: 17-MAR-1995  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: No. 5789389nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 95,332  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-405-702A-12

Query Match 0.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22  
Db 222 SLALVGA 228

## RESULT 60

US-08-690-095-4  
Sequence 4, Application US/08690095  
Patent No. 5792648  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Au-Young, Janice  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN MACROPHAGE ANTIGEN  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/690,095  
FILING DATE: Filed Herewith  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0110 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 179367  
US-08-690-095-4

Query Match 0.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 222 SLALVGA 228

RESULT 61  
US-08-465-485A-21  
Sequence 21, Application US/08465485A  
Patent No. 5831066  
GENERAL INFORMATION:  
APPLICANT: Reed, John  
TITLE OF INVENTION: Regulation of bcl-2 Gene Expression  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. Jefferson Davis Hwy., Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,485A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/124,256  
FILING DATE: 20-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/840,716  
FILING DATE: 21-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/288,692  
FILING DATE: 22-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Fortney, Andrew D.

REGISTRATION NUMBER: 34,600  
REFERENCE/DOCKET NUMBER: 3335-070-55 CONT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (408) 436-2070  
TELEFAX: (408) 436-2075  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-465-485A-21

Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 222 SLALVGA 228

RESULT 62  
US-08-661-479-51  
Sequence 51, Application US/08661479  
Patent No. 5834209  
GENERAL INFORMATION:  
APPLICANT: KORSMEYER, Stanley J.  
TITLE OF INVENTION: Bcl-x/Bcl-2 ASSOCIATED CELL DEATH  
TITLE OF INVENTION: REGULATOR  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Khourie and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/661,479  
FILING DATE: 11-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/333,565  
FILING DATE: 31-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15726A-000700  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-661-479-51

Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22

Db 222 SLALVGA 228  
|||||  
RESULT 63  
US-08-365-486A-15  
; Sequence 15, Application US/08365486A  
; Patent No. 5834306  
; GENERAL INFORMATION:  
; APPLICANT: Webster, Keith A.  
; APPLICANT: Bishopric, Nanette H.  
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
; TITLE OF INVENTION: Therapeutic Constructs  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/365,486A  
; FILING DATE: 23-DEC-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 8255-0018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-365-486A-15  
Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
QY 16 SLALVGA 22  
|||||  
Db 222 SLALVGA 228  
|||||  
RESULT 64  
US-08-365-486A-17  
; Sequence 17, Application US/08365486A  
; Patent No. 5834306  
; GENERAL INFORMATION:  
; APPLICANT: Webster, Keith A.  
; APPLICANT: Bishopric, Nanette H.  
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
; TITLE OF INVENTION: Therapeutic Constructs  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/365,486A  
; FILING DATE: 23-DEC-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 8255-0018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-365-486A-17  
Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
QY 16 SLALVGA 22  
|||||  
Db 222 SLALVGA 228  
|||||  
RESULT 65  
US-08-337-646A-10  
; Sequence 10, Application US/08337646A  
; Patent No. 5856171  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Hourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/337,646A  
; FILING DATE: 10-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/248,819  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/112,208  
; FILING DATE: 26-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000620  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids



TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-337-646A-10

Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
|||||||  
DB 222 SLALVGA 228

## RESULT 66

US-08-337-646A-12  
Sequence 12, Application US/08337646A  
Patent No. 5856171  
GENERAL INFORMATION:  
APPLICANT: KORSMEYER, Stanley J.  
TITLE OF INVENTION: CELL DEATH REGULATORS  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/337,646A  
FILING DATE: 10-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,819.  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/112,208  
FILING DATE: 26-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15726A-000620  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-337-646A-12

Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
|||||||  
DB 222 SLALVGA 228

## RESULT 67

US-08-408-095-16  
Sequence 16, Application US/08408095  
Patent No. 5858678  
GENERAL INFORMATION:  
APPLICANT: Chinnadurai, Govindaswamy  
TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/408,095  
FILING DATE: 21-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mack, Susan J.  
REGISTRATION NUMBER: 30,951  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-408-095-16

Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
|||||||  
DB 222 SLALVGA 228

## RESULT 68

US-08-856-531-10  
Sequence 10, Application US/08856531  
Patent No. 5942490  
GENERAL INFORMATION:  
APPLICANT: KORSMEYER, Stanley J.  
TITLE OF INVENTION: CELL DEATH REGULATORS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howell & Haferkamp, L.C.  
STREET: 7733 Forsyth Blvd., Suite 1400  
CITY: St. Louis  
STATE: MO  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/856,531  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: HOLLAND, Donald R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 976176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 314-727-5188  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..239  
OTHER INFORMATION: /note= "Human Bcl-2 polypeptide"  
US-08-856-034-10

Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 222 SLALVGA 228  
|||||

RESULT 69  
US-08-856-034-10  
; Sequence 10, Application US/08856034  
; Patent No. 5955595  
; GENERAL INFORMATION:  
; APPLICANT: KORMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howell & Haferkamp, L.C.  
; STREET: 7733 Forsyth Blvd., Suite 1400  
; CITY: St. Louis  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/856,034  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, Donald R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 976175  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..239  
; OTHER INFORMATION: /note= "Human Bcl-2 polypeptide"  
US-08-856-034-10

Query Match 0.5%; Score 7; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 222 SLALVGA 228  
|||||

RESULT 70  
US-08-471-057-12  
; Sequence 12, Application US/08471057  
; Patent No. 6015687  
; GENERAL INFORMATION:  
; APPLICANT: KIEFER, MICHAEL C.  
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,057  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/320,157  
; FILING DATE: 07-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LEHNHARDT, SUSAN K.  
; REGISTRATION NUMBER: 33,943  
; REFERENCE/DOCKET NUMBER: 23647-20007.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600  
; TELEFAX: (415) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-471-057-12

Query Match 0.5%; Score 7; DB 3; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 222 SLALVGA 228  
|||||

RESULT 71  
US-09-113-789-4  
; Sequence 4, Application US/09113789  
; Patent No. 6034219  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Au-Young, Janice.  
; APPLICANT: Goli, Surya K.

```

; TITLE OF INVENTION: NOVEL HUMAN MACROPHAGE ANTIGEN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,789
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/690,095
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0110 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 179367
; US-09-113-789-4

```

```

Query Match 0.5%; Score 7; DB 3; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 SLALVGA 22
|
Db 222 SLALVGA 228

```

```

RESULT 72
US-09-080-285-21
; Sequence 21, Application US/09080285
; Patent No. 6040181
; GENERAL INFORMATION:
; APPLICANT: Reed, John
; TITLE OF INVENTION: Regulation of bcl-2 Gene Expression
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 S. Jefferson Davis Hwy., Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,285
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,485
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/124,256
; FILING DATE: 20-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/840,716
; FILING DATE: 21-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/288,692
; FILING DATE: 22-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fortney, Andrew D.
; REGISTRATION NUMBER: 34,600
; REFERENCE/DOCKET NUMBER: 3335-070-55 CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 436-2070
; TELEFAX: (408) 436-2075
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-080-285-21

```

```

Query Match 0.5%; Score 7; DB 3; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 SLALVGA 22
|
Db 222 SLALVGA 228

```

```

RESULT 73
US-09-127-048-8
; Sequence 8, Application US/09127048
; Patent No. 6165732
; GENERAL INFORMATION:
; APPLICANT: Korsmeyer, Stanley J.
; APPLICANT: Schlesinger, Paul H.
; TITLE OF INVENTION: Method for Identifying Apoptosis Modulating Compounds
; FILE REFERENCE: 6029-6052
; CURRENT APPLICATION NUMBER: US/09/127,048
; CURRENT FILING DATE: 1998-07-31
; EARLIER APPLICATION NUMBER: 60/061,823
; EARLIER FILING DATE: 1997-10-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-127-048-8

```

```

Query Match 0.5%; Score 7; DB 4; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 SLALVGA 22
|
Db 222 SLALVGA 228

```

```

RESULT 74
US-08-927-326-10
; Sequence 10, Application US/08927326
; Patent No. 6184202

```

```
;; GENERAL INFORMATION:
;; APPLICANT: KORSMEYER, Stanley J.
;; TITLE OF INVENTION: CELL DEATH REGULATORS
;; NUMBER OF SEQUENCES: 78
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend Hourie and Crew
;; STREET: 379 Lytton Avenue
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: US
;; ZIP: 94301
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/927,326
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/337,646
;; FILING DATE: 10-NOV-1994
;; APPLICATION NUMBER: US 08/248,819
;; FILING DATE: 25-MAY-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/112,208
;; FILING DATE: 26-AUG-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, William M
;; REGISTRATION NUMBER: 30,223
;; REFERENCE/DOCKET NUMBER: 15726A-000620
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 239 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-927-326-10

Query Match 0.5%; Score 7; DB 4; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22
Db 222 SLALVGA 228

RESULT 75
US-08-927-326-12
; Sequence 12, Application US/08927326
; Patent No. 6184202
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, Stanley J.
; TITLE OF INVENTION: CELL DEATH REGULATORS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Hourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/927,326
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/337,646
;; FILING DATE: 10-NOV-1994
;; APPLICATION NUMBER: US 08/248,819
;; FILING DATE: 25-MAY-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/112,208
;; FILING DATE: 26-AUG-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, William M
;; REGISTRATION NUMBER: 30,223
;; REFERENCE/DOCKET NUMBER: 15726A-000620
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 239 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-927-326-12

Query Match 0.5%; Score 7; DB 4; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22
Db 222 SLALVGA 228

RESULT 76
US-08-880-342-15
; Sequence 15, Application US/08880342
; Patent No. 6218179
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; APPLICANT: Murphy, Brian
; APPLICANT: Laderoute, Keith R.
; APPLICANT: Green, Christopher J.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,342
; FILING DATE: 23-JUN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB95/00996
; FILING DATE: 13-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/365,486
```

```
;; FILING DATE: 23-DEC-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sholtz, Charles K.
;; REGISTRATION NUMBER: 38,615
;; REFERENCE/DOCKET NUMBER: 8255-0018.30
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 324-0880
;; TELEFAX: (415) 324-0960
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 239 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-880-342-15

Query Match 0.5%; Score 7; DB 4; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22
Db 222 SLALVGA 228

RESULT 77
US-08-880-342-17
Sequence 17, Application US/08880342
Patent No. 6218179
GENERAL INFORMATION:
APPLICANT: Webster, Keith A.
APPLICANT: Bishopric, Nanette H.
APPLICANT: Murphy, Brian
APPLICANT: Laderoute, Keith R.
APPLICANT: Green, Christopher J.
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
TITLE OF INVENTION: Therapeutic Constructs
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,342
FILING DATE: 23-JUN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB95/00996
FILING DATE: 13-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/365,486
FILING DATE: 23-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
TOPOLOGY: linear

;; FILING DATE: 23-DEC-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sholtz, Charles K.
;; REGISTRATION NUMBER: 38,615
;; REFERENCE/DOCKET NUMBER: 8255-0018.30
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 324-0880
;; TELEFAX: (415) 324-0960
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 239 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-880-342-17

Query Match 0.5%; Score 7; DB 4; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22
Db 222 SLALVGA 228

RESULT 78
PCT-US93-05651-5
Sequence 5, Application PC/TUS9305651
GENERAL INFORMATION:
TITLE OF INVENTION: A Gene Which Prevents Programmed Cell Death
NUMBER OF SEQUENCES: 5
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05651
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-05651-5

Query Match 0.5%; Score 7; DB 5; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22
Db 222 SLALVGA 228

RESULT 79
PCT-US95-04600-20
Sequence 20, Application PC/TUS9504600
GENERAL INFORMATION:
APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION
TITLE OF INVENTION: Interaction of Proteins Involved in
TITLE OF INVENTION: a Cell Death Pathway
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04600
FILING DATE: 12-APR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Imbra, Richard J.
REGISTRATION NUMBER: 37,643
REFERENCE/DOCKET NUMBER: FP-LJ 1361
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-8949
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 20:
```

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
PCT-US95-04600-20

Query Match 0.5%; Score 7; DB 5; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
| | | | |  
Db 222 SLALVGA 228

RESULT 80  
US-08-463-903-24  
; Sequence 24, Application US/08463903  
; Patent No. 6071515  
; GENERAL INFORMATION:  
; APPLICANT: Mezes, Peter S.  
; APPLICANT: Richard, Ruth A.  
; APPLICANT: Afholter, Joseph A.  
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides  
; FILE REFERENCE: 40224A US  
; CURRENT APPLICATION NUMBER: US/08/463,903  
; CURRENT FILING DATE: 1995-06-05  
; EARLIER APPLICATION NUMBER: US 07/935,695  
; EARLIER FILING DATE: 1992-08-21  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: MS-Word for Windows, Ver. 7.0  
; SEQ ID NO: 24  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: single chain pel B peptide-HLA-DRb11/a-FLAG construct  
; LOCATION: 1..244  
US-08-463-903-24

Query Match 0.5%; Score 7; DB 6; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
| | | | |  
Db 222 SLALVGA 228

RESULT 81  
US-08-463-903-26  
; Sequence 26, Application US/08463903  
; Patent No. 6071515  
; GENERAL INFORMATION:  
; APPLICANT: Mezes, Peter S.  
; APPLICANT: Richard, Ruth A.  
; APPLICANT: Afholter, Joseph A.  
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR BCL-2 GENE PRODUCT  
; NUMBER OF SEQUENCES: 5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/435,193  
; FILING DATE: 05-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 228,704  
; FILING DATE: 18-APR-1994  
; APPLICATION NUMBER: 994,941  
; FILING DATE: 23-DEC-1992  
; APPLICATION NUMBER: 663,010  
; FILING DATE: 19-MAR-1991  
; APPLICATION NUMBER: 883,687  
; FILING DATE: 09-JUL-1986  
; SEQ ID NO: 2  
; LENGTH: 239  
5506344-2

Query Match 0.5%; Score 7; DB 6; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
| | | | |  
Db 222 SLALVGA 228

RESULT 82  
US-08-463-903-24  
; Sequence 24, Application US/08463903  
; Patent No. 6071515  
; GENERAL INFORMATION:  
; APPLICANT: Mezes, Peter S.  
; APPLICANT: Richard, Ruth A.  
; APPLICANT: Afholter, Joseph A.  
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides  
; FILE REFERENCE: 40224A US  
; CURRENT APPLICATION NUMBER: US/08/463,903  
; CURRENT FILING DATE: 1995-06-05  
; EARLIER APPLICATION NUMBER: US 07/935,695  
; EARLIER FILING DATE: 1992-08-21  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: MS-Word for Windows, Ver. 7.0  
; SEQ ID NO: 24  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: single chain pel B peptide-HLA-DRb11/a-FLAG construct  
; LOCATION: 1..244  
US-08-463-903-24

Query Match 0.5%; Score 7; DB 3; Length 244;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
| | | | |  
Db 204 FEAQAL 210

RESULT 83  
US-08-463-903-26  
; Sequence 26, Application US/08463903  
; Patent No. 6071515  
; GENERAL INFORMATION:  
; APPLICANT: Mezes, Peter S.  
; APPLICANT: Richard, Ruth A.  
; APPLICANT: Afholter, Joseph A.  
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides  
; FILE REFERENCE: 40224A US  
; CURRENT APPLICATION NUMBER: US/08/463,903  
; CURRENT FILING DATE: 1995-06-05  
; EARLIER APPLICATION NUMBER: US 07/935,695  
; EARLIER FILING DATE: 1992-08-21  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: MS-Word for Windows, Ver. 7.0  
; SEQ ID NO: 26  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: single chain pel B peptide-HLA-DRb42/a-FLAG construct  
; LOCATION: 1..244  
US-08-463-903-26

Query Match 0.5%; Score 7; DB 3; Length 244;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
|||||  
DB 204 FEAQAL 210

RESULT 84  
US-08-463-903-28  
; Sequence 28, Application US/08463903  
; Patent No. 6071515  
; GENERAL INFORMATION:  
; APPLICANT: Mezes, Peter S.  
; APPLICANT: Richard, Ruth A.  
; APPLICANT: Affholter, Joseph A.  
; APPLICANT: Kotite, Nicolas J.  
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides  
; FILE REFERENCE: 40224A US  
; CURRENT APPLICATION NUMBER: US/08/463,903  
; CURRENT FILING DATE: 1995-06-05  
; EARLIER APPLICATION NUMBER: US 07/935,695  
; EARLIER FILING DATE: 1992-08-21  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: MS-Word for Windows, Ver. 7.0  
; SEQ ID NO 28  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: single chain pel B peptide-HLA-DRb48/a-FLAG construct  
; LOCATION: 1..244  
US-08-463-903-28

Query Match 0.5%; Score 7; DB 3; Length 244;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1109 FEAQAL 1115  
|||||  
DB 204 FEAQAL 210

RESULT 85  
US-08-463-903-30  
; Sequence 30, Application US/08463903  
; Patent No. 6071515  
; GENERAL INFORMATION:  
; APPLICANT: Mezes, Peter S.  
; APPLICANT: Richard, Ruth A.  
; APPLICANT: Affholter, Joseph A.  
; APPLICANT: Kotite, Nicolas J.  
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides  
; FILE REFERENCE: 40224A US  
; CURRENT APPLICATION NUMBER: US/08/463,903  
; CURRENT FILING DATE: 1995-06-05  
; EARLIER APPLICATION NUMBER: US 07/935,695  
; EARLIER FILING DATE: 1992-08-21  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: MS-Word for Windows, Ver. 7.0  
; SEQ ID NO 30  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: single chain pel B peptide-HLA-DRb41/a-FLAG construct  
; LOCATION: 1..244  
US-08-463-903-30

Query Match 0.5%; Score 7; DB 3; Length 244;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1109 FEAQAL 1115

Db 204 FEAQAL 210

RESULT 86  
US-08-644-664B-27  
; Sequence 27, Application US/08644664B  
; Patent No. 5776746  
; GENERAL INFORMATION:  
; APPLICANT: Denney Jr., Dan W.  
; TITLE OF INVENTION: Gene Amplification Methods  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/644,664B  
; APPLICATION NUMBER: US/08/644,664B  
; FILING DATE: 01-MAY-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: GENITOP-00912  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 248 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-644-664B-27

Query Match 0.5%; Score 7; DB 1; Length 248;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
|||||  
DB 79 FEAQAL 85

RESULT 87  
US-08-761-277A-27  
; Sequence 27, Application US/08761277A  
; Patent No. 5972334  
; GENERAL INFORMATION:  
; APPLICANT: Denney Jr., Dan W.  
; TITLE OF INVENTION: Vaccines For Treatment Of Lymphoma And  
; TITLE OF INVENTION: Leukemia  
; NUMBER OF SEQUENCES: 80  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/761,277A  
FILING DATE: 06-DEC-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/644,664  
FILING DATE: 01-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MacKnight, Kamrin T.  
REGISTRATION NUMBER: 38,230  
REFERENCE/DOCKET NUMBER: GENITOPPE-02406  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 248 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-761-277A-27

Query Match 0.5% Score 7; DB 2; Length 248;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
|||||  
Db 79 FEAQAL 85

RESULT 88  
US-08-484-905-109  
Sequence 109, Application US/08484905  
Patent No. 5976551  
GENERAL INFORMATION:  
APPLICANT: Mottez, Estelle  
APPLICANT: Abastado, Jean-Pierre  
APPLICANT: Kourilsky, Philippe  
TITLE OF INVENTION: An Altered Major Histocompatibility  
TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the  
TITLE OF INVENTION: Determinant  
NUMBER OF SEQUENCES: 127  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,905  
FILING DATE: 07-JUNE-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/801,818  
FILING DATE: 05-DEC-1991  
CLASSIFICATION: 530  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/792,473  
FILING DATE: 15-NOV-1991  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 03495.0106-03000

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 109:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Region  
LOCATION: 7  
OTHER INFORMATION: /note= "Xaa is an unidentified  
OTHER INFORMATION: amino acid residue."  
US-08-484-905-109  
Query Match 0.5% Score 7; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1109 FEAQAL 1115  
|||||  
Db 80 FEAQAL 86  
RESULT 89  
US-08-481-985B-109  
Sequence 109, Application US/08481985B  
Patent No. 6011146  
GENERAL INFORMATION:  
APPLICANT: Mottez, Estelle  
APPLICANT: Abastado, Jean-Pierre  
APPLICANT: Kourilsky, Philippe  
TITLE OF INVENTION: Altered Major Histocompatibility Complex  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,985B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/801,818  
FILING DATE: 05-DEC-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/792,473  
FILING DATE: 15-NOV-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03495.0106-04000  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 109:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear



MOLECULE TYPE: peptide  
FEATURE:

NAME/KEY: Region  
LOCATION: 7

OTHER INFORMATION: /note= "Xaa is an unidentified amino acid residue."

US-08-481-985B-109

## Query Match

Best Local Similarity 0.58; Score 7; DB 3; Length 253;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115

Db 80 FEAQAL 86

## RESULT 90

US-08-370-476-109

Sequence 109, Application US/08370476

Patent No. 6153408

## GENERAL INFORMATION:

APPLICANT: Mottez, Estelle

APPLICANT: Abastado, Jean-Pierre

APPLICANT: Kourilsky, Philippe

APPLICANT: Lone, Yu-Chun

APPLICANT: Ojcius, David

APPLICANT: Casrouge, Amanda

TITLE OF INVENTION: Altered Major Histocompatibility Complex

NUMBER OF SEQUENCES: 127

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1300 I Street, N.W., Suite 700

CITY: Washington

STATE: D.C.

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/370,476

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/117,575

FILING DATE: 07-SEP-1993

APPLICATION NUMBER: US 08/072,787

FILING DATE: 06-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/801,818

FILING DATE: 05-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/792,473

FILING DATE: 15-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 05243.0001-01000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 109:

SEQUENCE CHARACTERISTICS:

LENGTH: 253 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Region

LOCATION: 7

OTHER INFORMATION: /note= "Xaa is an unidentified amino acid residue."

US-08-370-476-109

## Query Match

Best Local Similarity 0.58; Score 7; DB 4; Length 253;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115

Db 80 FEAQAL 86

## RESULT 91

US-08-484-905-114

Sequence 114, Application US/08484905

Patent No. 5976551

## GENERAL INFORMATION:

APPLICANT: Mottez, Estelle

APPLICANT: Abastado, Jean-Pierre

APPLICANT: Kourilsky, Philippe

TITLE OF INVENTION: An Altered Major Histocompatibility

TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the

TITLE OF INVENTION: Determinant

NUMBER OF SEQUENCES: 127

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1300 I Street, N.W., Suite 700

CITY: Washington

STATE: D.C.

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS-/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,905

FILING DATE: 07-JUNE-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/801,818

FILING DATE: 05-DEC-1991

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/792,473

FILING DATE: 15-NOV-1991

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E. R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 03495.0106-03000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 114:

SEQUENCE CHARACTERISTICS:

LENGTH: 256 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Region

LOCATION: 7

OTHER INFORMATION: /note= "Xaa is an unidentified amino acid residue."

US-08-484-905-114

Query Match 0.5%; Score 7; DB 2; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAGAL 1115  
|||||  
DB 80 FEAGAL 86

## RESULT 92

US-08-481-985B-114  
; Sequence 114, Application US/08481985B  
; Patent No. 6011146  
; GENERAL INFORMATION:  
; APPLICANT: Mottez, Estelle  
; APPLICANT: Abastado, Jean-Pierre  
; APPLICANT: Kourilsky, Philippe  
; TITLE OF INVENTION: Altered Major Histocompatibility Complex  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/481,985B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/801,818  
; FILING DATE: 05-DEC-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/792,473  
; FILING DATE: 15-NOV-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0106-04000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 256 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 7  
; OTHER INFORMATION: /note= "Xaa is an unidentified  
; OTHER INFORMATION: amino acid residue."  
US-08-481-985B-114

Query Match 0.5%; Score 7; DB 3; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAGAL 1115  
|||||  
DB 80 FEAGAL 86

## RESULT 93

US-08-370-476-114  
; Sequence 114, Application US/08370476  
; Patent No. 6153408  
; GENERAL INFORMATION:  
; APPLICANT: Mottez, Estelle  
; APPLICANT: Abastado, Jean-Pierre  
; APPLICANT: Kourilsky, Philippe  
; APPLICANT: Lone, Yu-Chun  
; APPLICANT: Ojcius, David  
; APPLICANT: Castrouge, Amanda  
; TITLE OF INVENTION: Altered Major Histocompatibility Complex  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 127  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370,476  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/117,575  
; FILING DATE: 07-SEP-1993  
; APPLICATION NUMBER: US 08/072,787  
; FILING DATE: 06-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/801,818  
; FILING DATE: 05-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/792,473  
; FILING DATE: 15-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 05243.0001-01000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 256 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 7  
; OTHER INFORMATION: /note= "Xaa is an unidentified  
; OTHER INFORMATION: amino acid residue."  
US-08-370-476-114

Query Match 0.5%; Score 7; DB 4; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAGAL 1115  
|||||  
DB 80 FEAGAL 86

## RESULT 94

US-08-207-481-39  
; Sequence 39, Application US/08207481  
; Patent No. 5820866  
; GENERAL INFORMATION:  
; APPLICANT: Kappler, John W.  
; APPLICANT: Marrack, Philippa  
; TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL  
; TITLE OF INVENTION: REGULATION  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
; STREET: 1700 LINCOLN STREET, SUITE 3500  
; CITY: DENVER  
; STATE: COLORADO  
; COUNTRY: USA  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/207,481  
; FILING DATE: 04-MAR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kovarik, Joseph E.  
; REGISTRATION NUMBER: 33,005  
; REFERENCE/DOCKET NUMBER: 2879-8  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-207-481-39

Query Match 0.5%; Score 7; DB 2; Length 298;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1109 FEAOGAL 1115  
Db 34 FEAOGAL 40  
  
RESULT 95  
PCT-US95-02689-41  
; Sequence 41, Application PC/TUS9502689  
; GENERAL INFORMATION:  
; APPLICANT: National Jewish Center for Immunology and  
; APPLICANT: Respiratory Medicine  
; APPLICANT: Kappler, John W.  
; APPLICANT: Marrack, Philippa  
; TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL REGULATION  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
; STREET: 1700 LINCOLN STREET, SUITE 3500  
; CITY: DENVER  
; STATE: COLORADO  
; COUNTRY: USA  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/02689  
; FILING DATE: 03-MAR-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kovarik, Joseph E.  
; REGISTRATION NUMBER: 33,005  
; REFERENCE/DOCKET NUMBER: 2879-8-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 41:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-02689-41

Query Match 0.5%; Score 7; DB 5; Length 298;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1109 FEAOGAL 1115  
Db 34 FEAOGAL 40  
  
RESULT 96  
US-08-904-234-1  
; Sequence 1, Application US/08904234  
; Patent No. 6232459  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Tang, Tom Y.  
; TITLE OF INVENTION: NEW SYNAPTOTANIN ISOFORM  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/904,234  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0357 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 305 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: SYNORAT01  
; CLONE: 367401

US-08-904-234-1

Query Match 0.5%; Score 7; DB 4; Length 305;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VKTNGIS 566  
| | | | |  
DB 36 VKTNGIS 42

RESULT 97

US-08-871-268A-18  
; Sequence 18, Application US/08871268A  
; Patent No. 5866391  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Aubrey  
; APPLICANT: Cherry, Joel R.  
; TITLE OF INVENTION: Aspergillus Porphobilinogen Synthases  
; TITLE OF INVENTION: and Nucleic Acids Encoding Same  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 58663910 No. 5866391disk of No. 5866391th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871,268A  
; FILING DATE: 09-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4809.200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 330 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 5866391e  
US-08-871-268A-18

Query Match 0.5%; Score 7; DB 2; Length 330;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 TATTTLN 954  
| | | | |  
DB 21 TATTTLN 27

RESULT 98

US-08-871-267B-26  
; Sequence 26, Application US/08871267B  
; Patent No. 6100057  
; GENERAL INFORMATION:  
; APPLICANT: Elrod, Susan L.  
; APPLICANT: Cherry, Joel R.  
; APPLICANT: Jones, Aubrey  
; TITLE OF INVENTION: A Method for Increasing Hemoprotein

; TITLE OF INVENTION: Production in Filamentous Fungi  
; NUMBER OF SEQUENCES: 39

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 61000570 No. 6100057disk Of No. 6100057th America, Inc.  
; STREET: 405 Lexington Avenue - 64th Fl.  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871,267B  
; FILING DATE: 9-JUN-1997  
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: Rozek, Carol E.  
; REGISTRATION NUMBER: 36,993  
; REFERENCE/DOCKET NUMBER: 4771.200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-878-9652  
; TELEFAX: 212-878-9655  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6100057e  
US-08-871-267B-26

; INFORMATION FOR SEQ ID NO: 26:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 330 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6100057e  
US-08-871-267B-26

Query Match 0.5%; Score 7; DB 3; Length 330;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 TATTTLN 954  
| | | | |  
DB 21 TATTTLN 27

RESULT 99

US-08-933-750C-26  
; Sequence 26, Application US/08933750C  
; Patent No. 5932442

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Bandman, Olga  
; APPLICANT: Shah, Purvi  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Yue, Henry  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
; NUMBER OF SEQUENCES: 98  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/933,750C  
FILING DATE: September 23, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LNOBNOT03  
CLONE: 1573677  
US-08-933-750C-26

Query Match 0.5%; Score 7; DB 2; Length 340;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 YGTSAGV 1055  
|||||||  
DB 230 YGTSAGV 236

RESULT 100  
US-09-234-613-26  
Sequence 26, Application US/09234613  
Patent No. 6132973  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/234,613  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750  
FILING DATE: September 23, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LNOBNOT03  
CLONE: 1573677  
US-09-234-613-26

Query Match 0.5%; Score 7; DB 4; Length 340;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 YGTSAGV 1055  
|||||||  
DB 230 YGTSAGV 236

RESULT 101  
US-08-103-170-8  
Sequence 8, Application US/08103170  
Patent No. 5885824  
GENERAL INFORMATION:  
APPLICANT: Yamada, Tadataka  
APPLICANT: Gantz, Ira  
TITLE OF INVENTION: Recombinant Genomic Clones Encoding  
TITLE OF INVENTION: Histamine H1, H2, and H3 Receptors, Methods For Production  
TITLE OF INVENTION: Thereof, and Proteins Encoded Therefrom  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/103,170  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/633,060  
FILING DATE: 24-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-017-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 24885 OPAT UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Hamster  
US-08-103-170-8

Query Match 0.5%; Score 7; DB 2; Length 419;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1065 AIVGGFG 1071  
Db 45 AIVGGFG 51

## RESULT 102

US-08-705-771-13  
; Sequence 13, Application US/08705771  
; Patent No. 6054289  
; GENERAL INFORMATION:  
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
; APPLICANT: Jian Ni and Jing-Shan Hu  
; TITLE OF INVENTION: Human Genes, Sequences and  
; TITLE OF INVENTION: Expression Products  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1

## CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/705,771  
; FILING DATE: August 30, 1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:

; NAME: MULLINS, J.G.  
; REGISTRATION NUMBER: 33,073  
; REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 973-994-1700  
; TELEFAX: 973-994-1744

## INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 443 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-705-771-13

Query Match 0.5%; Score 7; DB 3; Length 443;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LKQAEA 71  
Db 131 LKQAEA 137

## RESULT 103

US-08-679-635A-4  
; Sequence 4, Application US/08679635A  
; Patent No. 5985643  
; GENERAL INFORMATION:

; APPLICANT: Tomas, Alexander  
; APPLICANT: Delencastre, Herminia  
; TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF  
; TITLE OF INVENTION: METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF  
; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; STREET: Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/679,635A  
; APPLICATION NUMBER: US/08/679,635A  
; FILING DATE: 10-JUL-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-141  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684

## INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 463 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO

US-08-679-635A-4

Query Match 0.5%; Score 7; DB 2; Length 463;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 TSAGVDA 1057  
Db 67 TSAGVDA 73

## RESULT 104

US-09-118-442-11  
; Sequence 11, Application US/09118442B  
; Patent No. 6197561  
; GENERAL INFORMATION:

; APPLICANT: Martino-Catt, Susan J.  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Beach, Larry R.  
; APPLICANT: Wang, Xun

; APPLICANT: Bowen, Benjamin A.

; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in  
; TITLE OF INVENTION: Plants and Uses Thereof

## FILE REFERENCE: 0706

; CURRENT APPLICATION NUMBER: US/09/118,442B  
; CURRENT FILING DATE: 1998-07-17  
; EARLIER APPLICATION NUMBER: 60/055,446  
; EARLIER FILING DATE: 1997-08-11  
; EARLIER APPLICATION NUMBER: 60/055,526  
; EARLIER FILING DATE: 1997-08-08  
; EARLIER APPLICATION NUMBER: 60/053,944  
; EARLIER FILING DATE: 1997-07-28  
; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 3.0

## SEQ ID NO 11

; LENGTH: 510  
; TYPE: PRT  
; ORGANISM: Zea mays

US-09-118-442-11

Query Match 0.5%; Score 7; DB 4; Length 510;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
Db 333 VSYNHLG 339

RESULT 105  
US-08-808-931-16  
; Sequence 16, Application US/08808931  
; Patent No. 5939602  
; GENERAL INFORMATION:  
; APPLICANT: Volrath, Sandra  
; APPLICANT: Johnson, Marie  
; APPLICANT: Potter, Sharon  
; APPLICANT: Ward, Eric  
; APPLICANT: Heifetz, Peter  
; TITLE OF INVENTION: DNA Molecules Encoding Plant  
; TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: No. 5939602artis Corporation  
; STREET: 520 White Plains Road, P.O. Box 2005  
; CITY: Tarrytown  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10591-9005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,931  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/012,705  
; FILING DATE: 28-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/013,612  
; FILING DATE: 28-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/020,003  
; FILING DATE: 21-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: CGC 1847  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919) 541-8587  
; TELEFAX: (919) 541-8689  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 539 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
US-08-808-931-16

Query Match 0.5%; Score 7; DB 2; Length 539;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112  
Db 48 SSKIDGG 54

RESULT 106  
US-08-808-323-16  
; Sequence 16, Application US/08808323  
; Patent No. 6018105  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Marie  
; APPLICANT: Volrath, Sandra  
; APPLICANT: Ward, Eric  
; TITLE OF INVENTION: Promoters from Plant  
; TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6018105artis Corporation  
; STREET: 520 White Plains Road, P.O. Box 2005  
; CITY: Tarrytown  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10591-9005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,323  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/012,705  
; FILING DATE: 28-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/013,612  
; FILING DATE: 28-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/020,003  
; FILING DATE: 21-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: CGC 1846  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919) 541-8587  
; TELEFAX: (919) 541-8689  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 539 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
US-08-808-323-16

Query Match 0.5%; Score 7; DB 3; Length 539;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112  
Db 48 SSKIDGG 54

RESULT 107  
US-09-050-603A-16  
; Sequence 16, Application US/09050603A  
; Patent No. 6023012  
; GENERAL INFORMATION:  
; APPLICANT: Volrath, Sandra  
; APPLICANT: Johnson, Marie  
; APPLICANT: Potter, Sharon  
; APPLICANT: Ward, Eric

APPLICANT: Heifetz, Peter  
TITLE OF INVENTION: DNA Molecules Encoding Plant  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6023012artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 09/050,603A  
FILING DATE: 30-MAR-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/808,931  
FILING DATE: 28-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/012,705  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/013,612  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/020,003  
FILING DATE: 21-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1847  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8587  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 539 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-09-050-603A-16

Query Match 0.5%; Score 7; DB 3; Length 539;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112  
Db 48 SSKIDGG 54

RESULT 108  
US-09-102-420B-16  
Sequence 16, Application US/09102420B  
Patent No. 6084155  
GENERAL INFORMATION:  
APPLICANT: Volrath, Sandra  
APPLICANT: Johnson, Marie  
APPLICANT: Ward, Eric  
APPLICANT: Heifetz, Peter  
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN  
TITLE OF INVENTION: OXIDASE ("PROTOX")  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6084155artis Corporation  
STREET: 3054 Cornwallis Road

CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/102,420B  
FILING DATE: 22-JUN-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/059,164  
FILING DATE: 13-APR-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/050,603  
FILING DATE: 30-MAR-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/126,430  
FILING DATE: 11-MAR-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/808,931  
FILING DATE: 28-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/012,705  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/013,612  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/020,003  
FILING DATE: 21-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/472,028  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8587  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 539 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-09-102-420B-16

Query Match 0.5%; Score 7; DB 3; Length 539;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112  
Db 48 SSKIDGG 54

RESULT 109  
US-08-136-743B-4  
Sequence 4, Application US/08136743B  
Patent No. 5459063  
GENERAL INFORMATION:  
APPLICANT: Barry S. Cooperman, Harvey Rubin,  
APPLICANT: Jerome Salem, and Alison L. Fisher  
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-  
TITLE OF INVENTION: cletotide Reductase, DNA Sequences Therefor and Peptide Inhi  
TITLE OF INVENTION: Thereof"



NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The University of Pennsylvania  
STREET: Suite 330  
STREET: 3700 Market Street  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19104-3246  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/136,743B  
FILING DATE: 10/14/93  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 3957-10  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5459063e  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 811 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-136-743B-4

Query Match 0.5%; Score 7; DB 1; Length 811;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 750 GTNGISN 756  
Db 263 GTNGISN 269

RESULT 110  
US-08-457-176-2  
Sequence 2, Application US/08457176  
Patent No. 5591826  
GENERAL INFORMATION:  
APPLICANT: Vogelstein, Bert  
APPLICANT: Kinzler, Kenneth W.  
APPLICANT: de la Chappelle, Albert  
TITLE OF INVENTION: Mutator Gene and Hereditary  
TITLE OF INVENTION: No. 5591826-polyposis Colorectal Cancer  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie, and Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,176  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION NUMBER: US 08/160295  
FILING DATE: 02-DEC-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.44900  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202.508.9100  
TELEFAX: 202.508.9299  
TELEX: 197430 BBMB UT  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 934 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-457-176-2

Query Match 0.5%; Score 7; DB 1; Length 934;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1010 LESAAEV 1016  
Db 11 LESAAEV 17

RESULT 111  
US-08-457-175-2  
Sequence 2, Application US/08457175  
Patent No. 5693470  
GENERAL INFORMATION:  
APPLICANT: Vogelstein, Bert  
APPLICANT: Kinzler, Kenneth W.  
APPLICANT: de la Chappelle, Albert  
TITLE OF INVENTION: Mutator Gene and Hereditary  
TITLE OF INVENTION: No. 5693470-Polyposis Colorectal Cancer  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie, and Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,175  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US 08/160295  
FILING DATE: 02-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.44900  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202.508.9100  
TELEFAX: 202.508.9299  
TELEX: 197430 BBMB UT  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 934 amino acids  
TYPE: amino acid

;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: YES  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapiens  
US-08-457-175-2

Query Match 0.5%; Score 7; DB 1; Length 934;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1010 LESAAEV 1016  
Db 11 LESAAEV 17

RESULT 112  
US-08-709-784-1  
; Sequence 1, Application US/08709784  
; Patent No. 6048701  
; GENERAL INFORMATION:  
; APPLICANT: The Johns Hopkins University  
; TITLE OF INVENTION: Antibody Detection of Mismatch Repair  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Allegretti, Ltd.  
; STREET: 1001 G Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20001-4597  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/709,784  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/480,351  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.57434  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 934 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
US-08-709-784-1

Query Match 0.5%; Score 7; DB 3; Length 934;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1010 LESAAEV 1016

Db 11 LESAAEV 17

RESULT 113  
US-08-446-137B-2  
; Sequence 2, Application US/08446137B  
; Patent No. 6162903  
; GENERAL INFORMATION:  
; APPLICANT: Trowern, Angus R.  
; APPLICANT: Atkinson, Anthony  
; APPLICANT: Murphy, Jonathan P.  
; APPLICANT: Laurence, Oliver S.  
; APPLICANT: Dugdaley, Clive J.  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
; TITLE OF INVENTION: FROM L PROTEIN AND THEIR USES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,137B  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.406  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1027 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-446-137B-2

Query Match 0.5%; Score 7; DB 4; Length 1027;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 707 NDIDSAT 713  
Db 133 NDIDSAT 139

RESULT 114  
US-09-413-814-42  
; Sequence 42, Application US/09413814  
; Patent No. 6225064  
; GENERAL INFORMATION:  
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH  
; APPLICANT: Bristol-Myers Squibb, Co.  
; APPLICANT: Beyer, Stefan  
; APPLICANT: Bloecker, Helmut  
; APPLICANT: Brandt, Petra  
; APPLICANT: Cino, Paul M  
; APPLICANT: Dougherty, Brian A  
; APPLICANT: Goldberg, Steven L  
; APPLICANT: Hofle, Gerhard  
; APPLICANT: Mueller, Joachim

APPLICANT: Reichenbach, Hans  
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or  
FILE REFERENCE: PCT/US 99/23535  
CURRENT APPLICATION NUMBER: US/09/413,814  
CURRENT FILING DATE: 1999-10-07  
EARLIER APPLICATION NUMBER: DE 198 46 493.2  
EARLIER FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 42  
LENGTH: 2539  
TYPE: PRT  
ORGANISM: Sorangium cellulosum  
US-09-413-814-42

Query Match 0.5%; Score 7; DB 4; Length 2539;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 LLSWGLK 66  
Db 1092 LLSWGLK 1098  
|||||

RESULT 115  
US-08-945-168-101  
Sequence 101, Application US/08945168  
Patent No. 5989548  
GENERAL INFORMATION:  
APPLICANT: DILLNER, JOAKIM  
TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA  
TITLE OF INVENTION: VIRUS  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/945,168  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE96/00533  
FILING DATE: 23-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9501512-9  
FILING DATE: 24-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-945-168-101

Query Match 0.5%; Score 6; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 545 STNVAV 550  
Db 2 STNVAV 7  
|||||

RESULT 116  
US-08-633-772-25  
Sequence 25, Application US/08633772  
Patent No. 5731155  
GENERAL INFORMATION:  
APPLICANT: Schreiber, Robert D.  
APPLICANT: Farrar, Michael A.  
APPLICANT: Greenlund, Andrew C.  
TITLE OF INVENTION: Compositions for Inhibition of  
TITLE OF INVENTION: Intracellular Transcription Factors and Methods  
TITLE OF INVENTION: Therefor  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howell & Haferkamp, LC  
STREET: 7733 Forsyth Boulevard  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63105

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/633,772  
FILING DATE: 21-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/141,499  
FILING DATE: 22-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Holland, Donald R.  
REGISTRATION NUMBER: 35197  
REFERENCE/DOCKET NUMBER: 964635  
TELEPHONE: 314-727-5188  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 5  
OTHER INFORMATION: /note= phosphorylated tyrosine  
US-08-633-772-25

Query Match 0.5%; Score 6; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 356 KQESSQ 361  
Db 7 KQESSQ 12  
|||||

RESULT 117  
PCT-US95-04121-45  
; Sequence 45, Application PC/TUS9504121  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Haptenated Peptides and Uses Thereof  
; NUMBER OF SEQUENCES: 62  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04121  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/222,206  
; FILING DATE: April 1, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vanstone, Darlene A.  
; REGISTRATION NUMBER: 35,279  
; REFERENCE/DOCKET NUMBER: 079.2PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6010  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
PCT-US95-04121-45

Query Match 0.5%; Score 6; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQA 1114  
Db 1 FEAQA 6

RESULT 118  
US-08-945-168-99  
; Sequence 99, Application US/08945168  
; Patent No. 5989548  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA  
; NUMBER OF SEQUENCES: 117  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/945,168  
; FILING DATE: 18-DEC-1997

CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/SE96/00533  
; FILING DATE: 23-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9501512-9  
; FILING DATE: 24-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 99:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-945-168-99

Query Match 0.5%; Score 6; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 STNVAV 550  
Db 5 STNVAV 10

RESULT 119  
PCT-US93-06751-71  
; Sequence 71, Application PC/TUS9306751  
; GENERAL INFORMATION:  
; APPLICANT: P. Keller, A. J. Conley, A.R. Shaw, B.A. Arnold  
; TITLE OF INVENTION: Immunological Conjugates of OMPc and  
; NUMBER OF SEQUENCES: 146  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merck & Co., Inc.  
; STREET: P.O. Box 2000  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/06751  
; FILING DATE: 19930719  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meredith, Roy D.  
; REGISTRATION NUMBER: 30,777  
; REFERENCE/DOCKET NUMBER: 18614  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 594-4678  
; TELEFAX: (908) 594-4720  
; TELEX: 138825  
; INFORMATION FOR SEQ ID NO: 71:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO

RESULT 121  
US-08-927-326-31  
; Sequence 31, Application US/08927326

GENERAL INFORMATION:  
 APPLICANT: Marsden, Howard S  
 TITLE OF INVENTION: PEPTIDE STRUCTURES AND THEIR USE IN  
 TITLE OF INVENTION: DIAGNOSIS OF HERPES SIMPLEX VIRUS TYPE 2  
 NUMBER OF SEQUENCES: 86  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nixon & Vanderhye PC  
 STREET: 8th Floor, 1100 No. 5965357th Glebe Road  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: USA  
 ZIP: 22201-4714  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/017,205  
FILING DATE: 02-FEB-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mitchard, Leonard C  
REGISTRATION NUMBER: 29,009  
REFERENCE/DOCKET NUMBER: 604-436  
TELEPHONE: (703)816-4000  
TELEFAX: (703)816-4100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide from HSV-2 glycoprotein G  
FRAGMENT TYPE: internal  
US-09-017-205-1

Query Match 0.5%; Score 6; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 INPPNS 374  
Db 10 INPPNS 15

RESULT 123  
US-08-945-168-100  
; Sequence 100, Application US/08945168  
; Patent No. 5989548  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA  
; VIRUS  
; NUMBER OF SEQUENCES: 117  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/945,168  
; FILING DATE: 18-DEC-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/SE96/00533  
; FILING DATE: 23-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9501512-9  
; FILING DATE: 24-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 100:

SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-945-168-100

Query Match 0.5%; Score 6; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 STNAV 550  
Db 3 STNAV 8

RESULT 124  
US-09-007-905-52  
; Sequence 52, Application US/09007905  
; Patent No. 6221585  
; GENERAL INFORMATION:  
; APPLICANT: Fran ois J.-M. Iris and Jean-Louis Pourny  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES UNDERLYING  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/007,905  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Antler, Adriane M.  
; REGISTRATION NUMBER: 32,605  
; REFERENCE/DOCKET NUMBER: 9408-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 52:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-09-007-905-52

Query Match 0.5%; Score 6; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1165 SVGVSY 1170  
Db 2 SVGVSY 7

RESULT 125  
US-08-945-168-98  
; Sequence 98, Application US/08945168

Patent No. 5989548  
GENERAL INFORMATION:  
APPLICANT: DILLNER, JOAKIM  
TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA  
VIRUS  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/945,168  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE96/00533  
FILING DATE: 23-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9501512-9  
FILING DATE: 24-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-945-168-98

Query Match 0.5%; Score 6; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 STNAV 550  
Db 5 STNAV 10

RESULT 126  
US-08-985-090-11  
Sequence 11, Application US/08985090  
Patent No. 5885893  
GENERAL INFORMATION:  
APPLICANT: Andrew D.J. Goodearl  
TITLE OF INVENTION: MUSCARINIC RECEPTORS AND USES THEREFOR  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/985,090  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Jean M. Silveri  
REGISTRATION NUMBER: 39,030  
REFERENCE/DOCKET NUMBER: MNI-032  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-985-090-11

Query Match 0.5%; Score 6; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 LITAST 546  
Db 3 LITAST 8

RESULT 127  
US-09-165-543-11  
Sequence 11, Application US/09165543  
Patent No. 6093545  
GENERAL INFORMATION:  
APPLICANT: Andrew D.J. Goodearl and Sandra Glucksman  
TITLE OF INVENTION: Muscarinic Receptors and Uses Therefor  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,543  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/042,780  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth A. Hanley  
REGISTRATION NUMBER: 33,505  
REFERENCE/DOCKET NUMBER: MNI-032CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

FRAGMENT TYPE: internal  
US-09-165-543-11

Query Match 0.5%; Score 6; DB 3; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 LITAST 546  
|||||  
DB 3 LITAST 8

RESULT 128  
US-09-165-543-18  
; Sequence 18, Application US/09165543  
; Patent No. 6093545  
; GENERAL INFORMATION:  
; APPLICANT: Andrew D.J. Goodearl and Sandra Glucksman  
; TITLE OF INVENTION: Muscarinic Receptors and Uses Therefor  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/165,543  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/042,780  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elizabeth A. Hanley  
; REGISTRATION NUMBER: 33,505  
; REFERENCE/DOCKET NUMBER: MNI-032CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)42-4214  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-09-165-543-18

Query Match 0.5%; Score 6; DB 3; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 LITAST 546  
|||||  
DB 3 LITAST 8

RESULT 129  
US-08-295-643-3  
; Sequence 3, Application US/08295643  
; Patent No. 5859219  
; GENERAL INFORMATION:  
; APPLICANT: COVER, TIMOTHY L.  
; APPLICANT: BLASER, MARTIN J.

; TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
; TITLE OF INVENTION: HELICOBACTER PYLORI AND METHODS TO USE SAME  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
; STREET: Suite 1200, 127 Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/295,643  
; FILING DATE: 26-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SPRATT, GWENDOLYN D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 2200.025  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 24 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-295-643-3

Query Match 0.5%; Score 6; DB 2; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AFFTTV 39  
|||||  
DB 1 AFFTTV 6

RESULT 130  
US-08-473-265-2  
; Sequence 2, Application US/08473265  
; Patent No. 6013463  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
; TITLE OF INVENTION: HELICOBACTER PYLORI AND METHODS TO USE THE SAME  
; FILE REFERENCE: 22000.0026  
; CURRENT APPLICATION NUMBER: US/08/473,265  
; CURRENT FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/284,747  
; EARLIER FILING DATE: 1994-08-02  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/No. 6013463e -  
; OTHER INFORMATION: synthetic construct  
US-08-473-265-2

Query Match 0.5%; Score 6; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;



Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AFTTV 39  
|11111|  
Db 1 AFTTV 6

## RESULT 131

US-08-284-747-2  
; Sequence 2, Application US/08284747A  
; Patent No. 6054132  
; GENERAL INFORMATION:  
; APPLICANT: Cover, Timothy L.  
; APPLICANT: Blaser, Martin J.  
; TITLE OF INVENTION: PURIFIED VACUOLATING TOXIN FROM  
; TITLE OF INVENTION: HELICOBACTER PYLORI AND METHODS TO USE THE SAME  
; FILE REFERENCE: 22000.0026  
; CURRENT APPLICATION NUMBER: US/08/284,747A  
; CURRENT FILING DATE: 1994-08-02  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: /No. 6054132e -  
; OTHER INFORMATION: Synthetic construct  
US-08-284-747-2

Query Match 0.5%; Score 6; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AFTTV 39  
|11111|  
Db 1 AFTTV 6

## RESULT 132

US-08-046-585-6  
; Sequence 6, Application US/08046585  
; Patent No. 5453362  
; GENERAL INFORMATION:  
; APPLICANT: Lamarco, Kelly  
; APPLICANT: Wilson, Angus  
; APPLICANT: Herr, Winship  
; TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:  
; TITLE OF INVENTION: HOST CELL FACTOR  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/046,585  
; FILING DATE: 12-APR-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-57503-1/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-046-585-6

Query Match 0.5%; Score 6; DB 1; Length 26;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 947 NTATT 952  
|11111|  
Db 18 NTATT 23

## RESULT 133

US-08-393-703-6  
; Sequence 6, Application US/08393703  
; Patent No. 5585239  
; GENERAL INFORMATION:  
; APPLICANT: Lamarco, Kelly  
; APPLICANT: Wilson, Angus  
; APPLICANT: Herr, Winship  
; TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:  
; TITLE OF INVENTION: HOST CELL FACTOR  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,703  
; FILING DATE: 24-FEB-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-57503-2/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-393-703-6

Query Match 0.5%; Score 6; DB 1; Length 26;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 947 NTATT 952  
|11111|  
Db 18 NTATT 23

RESULT 134  
PCT-US93-11721-6  
; Sequence 6, Application PCT/US9311721  
; GENERAL INFORMATION:  
; APPLICANT: Lamarco, Kelly  
; APPLICANT: Wilson, Angus  
; APPLICANT: Herr, Winship  
; TITLE OF INVENTION: A NOVEL EUKARYOTIC TRANSCRIPTION PROTEIN:  
; TITLE OF INVENTION: HOST CELL FACTOR  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/11721  
; FILING DATE: 03-DEC-1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: FP-57503-1/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US93-11721-6

Query Match 0.5%; Score 6; DB 5; Length 26;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 947 NTATTT 952  
Db 18 NTATTT 23

RESULT 135  
US-08-336-553A-18  
; Sequence 18, Application US/08336553A  
; Patent No. 6054264  
; GENERAL INFORMATION:  
; APPLICANT: CHIEN, DAVID Y.  
; APPLICANT: KUO, GEORGE  
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND  
; TITLE OF INVENTION: REAGENTS FOR USE THEREIN  
; NUMBER OF SEQUENCES: 75  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,553A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/060,400  
FILING DATE: 10-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: LEHNHARDT, SUSAN K.  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 22300-20947.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-336-553A-18

Query Match 0.5%; Score 6; DB 3; Length 29;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1015 EVLXQF 1020  
Db 11 EVLXQF 16

RESULT 136  
US-08-662-227-37  
; Sequence 37, Application US/08662227  
; Patent No. 5922320  
; GENERAL INFORMATION:  
; APPLICANT: VOGEL, CARL-WILHELM  
; APPLICANT: BREDEHORST, REINHORST  
; APPLICANT: KOCK, MICHAEL  
; APPLICANT: FRITZINGER, DAVID  
; TITLE OF INVENTION: RECOMBINANT PROCVF  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/662,227  
; FILING DATE: 14-JUN-1996  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 1126-0107-0X  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 37:

SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-662-227-37

Query Match 0.5%; Score 6; DB 2; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 769 LYNNNN 774  
|||||  
DB 25 LYNNNN 30

## RESULT 137

US-07-915-247A-21  
; Sequence 21, Application US/07915247A  
; Patent No. 5589452  
; GENERAL INFORMATION:

APPLICANT: Kristenansky, John L.  
APPLICANT: Nestor Jr., John J.  
APPLICANT: Ho, Teresa H.  
APPLICANT: Vickery, Brian H.  
APPLICANT: Bach, Chinh T.

TITLE OF INVENTION: ANALOGS OF PARATHYROID HORMONE AND  
TITLE OF INVENTION: PARATHYROID HORMONE RELATED PEPTIDE: SYNTHESIS AND USE  
TITLE OF INVENTION: FOR THE TREATMENT OF OSTEOPOROSIS  
NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
STREET: 3401 Hillview Ave.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/915,247A  
FILING DATE: 19920714  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 27610  
TELEPHONE: 415-855-6593  
TELEFAX: 415-496-3529

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: N-terminal

US-07-915-247A-21

Query Match 0.5%; Score 6; DB 1; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 RSLSS 107  
|||||  
DB 21 RSLSS 26

## RESULT 138

US-08-443-863-21  
; Sequence 21, Application US/08443863  
; Patent No. 5693616  
; GENERAL INFORMATION:

APPLICANT: Kristenansky, John L.  
APPLICANT: Nestor Jr., John J.  
APPLICANT: Ho, Teresa H.  
APPLICANT: Vickery, Brian H.  
APPLICANT: Bach, Chinh T.

TITLE OF INVENTION: ANALOGS OF PARATHYROID HORMONE AND  
TITLE OF INVENTION: PARATHYROID HORMONE RELATED PEPTIDE: SYNTHESIS AND USE  
TITLE OF INVENTION: FOR THE TREATMENT OF OSTEOPOROSIS  
NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
STREET: 3401 Hillview Ave.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/443,863  
FILING DATE: 14-JUL-1992  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 27610  
TELEPHONE: 415-855-6593  
TELEFAX: 415-496-3529

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: N-terminal

US-08-443-863-21

## Query Match

Best Local Similarity 100.0%; Score 6; DB 1; Length 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 RSLSS 107  
|||||

DB 21 RSLSS 26

## RESULT 139

US-08-448-070-21  
; Sequence 21, Application US/08448070  
; Patent No. 5695955  
; GENERAL INFORMATION:

APPLICANT: Kristenansky, John L.  
APPLICANT: Nestor Jr., John J.  
APPLICANT: Ho, Teresa H.  
APPLICANT: Vickery, Brian H.  
APPLICANT: Bach, Chinh T.

TITLE OF INVENTION: ANALOGS OF PARATHYROID HORMONE AND  
TITLE OF INVENTION: PARATHYROID HORMONE RELATED PEPTIDE: SYNTHESIS AND USE  
TITLE OF INVENTION: FOR THE TREATMENT OF OSTEOPOROSIS  
NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
;; STREET: 3401 Hillview Ave.  
;; CITY: Palo Alto  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94303  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM-PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/448,070  
;; FILING DATE: 14-JUL-1992  
;; CLASSIFICATION: 435  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Schmonsees, William  
;; REGISTRATION NUMBER: 31,796  
;; REFERENCE/DOCKET NUMBER: 27610  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-855-6593  
;; TELEFAX: 415-496-3529  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 34 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; FRAGMENT TYPE: N-terminal  
;;  
;; US-08-448-070-21

Query Match 0.5%; Score 6; DB 1; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 102 RSLSS 107  
Db 21 RSLSS 26

RESULT 140  
US-08-449-500-21  
; Sequence 21, Application US/08449500  
; Patent No. 5798225  
; GENERAL INFORMATION:  
; APPLICANT: Kristenasky, John L.  
; APPLICANT: Nestor Jr., John J.  
; APPLICANT: Ho, Teresa H.  
; APPLICANT: Vickery, Brian H.  
; APPLICANT: Bach, Chinh T.  
; TITLE OF INVENTION: ANALOGS OF PARATHYROID HORMONE AND  
; TITLE OF INVENTION: PARATHYROID HORMONE RELATED PEPTIDE: SYNTHESIS AND USE  
; FOR THE TREATMENT OF OSTEOPOROSIS  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
; STREET: 3401 Hillview Ave.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/449,500  
; FILING DATE: 18-JAN-1994  
; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Schmonsees, William  
;; REGISTRATION NUMBER: 31,796  
;; REFERENCE/DOCKET NUMBER: 27610-P1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-855-6593  
;; TELEFAX: 415-496-3529  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 34 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; FRAGMENT TYPE: N-terminal  
;;  
;; US-08-449-500-21

Query Match 0.5%; Score 6; DB 1; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 102 RSLSS 107  
Db 21 RSLSS 26

RESULT 141  
US-08-449-317A-21  
; Sequence 21, Application US/08449317A  
; Patent No. 5807823  
; GENERAL INFORMATION:  
; APPLICANT: Vickery, Brian H.  
; TITLE OF INVENTION: METHOD FOR TREATMENT OF CORTICOSTEROID  
; TITLE OF INVENTION: INDUCED OSTEOPENIA  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
; STREET: 3401 Hillview Ave.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/449,317A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schmonsees, William  
; REGISTRATION NUMBER: 31,796  
; REFERENCE/DOCKET NUMBER: 27610-P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-6593  
; TELEFAX: 415-496-3529  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
; US-08-449-317A-21

Query Match 0.5%; Score 6; DB 1; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 102 RSLSS 107  
Db 21 RSLSS 26

## RESULT 142

US-08-477-022-21  
; Sequence 21, Application US/08477022  
; Patent No. 5821225  
; GENERAL INFORMATION:  
; APPLICANT: Vickery, Brian H.  
; TITLE OF INVENTION: METHOD FOR TREATMENT OF CORTICOSTEROID  
; INDUCED OSTEOPENIA  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
; STREET: 3401 Hillview Ave.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,022  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schmonsees, William  
; REGISTRATION NUMBER: 31,796  
; REFERENCE/DOCKET NUMBER: 27610-P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-6593  
; TELEFAX: 415-496-3529  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
US-08-477-022-21

Query Match 0.5%; Score 6; DB 2; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 RSLSS 107  
Db 21 RSLSS 26

## RESULT 143

US-08-449-447-21  
; Sequence 21, Application US/08449447  
; Patent No. 5840837  
; GENERAL INFORMATION:  
; APPLICANT: Kristenansky, John L.  
; APPLICANT: Nestor Jr., John J.  
; APPLICANT: Ho, Teresa H.  
; APPLICANT: Vickery, Brian H.  
; APPLICANT: Bach, Chinh T.  
; TITLE OF INVENTION: ANALOGS OF PARATHYROID HORMONE AND  
; PARATHYROID HORMONE RELATED PEPTIDE: SYNTHESIS AND USE  
; FOR THE TREATMENT OF OSTEOPOROSIS  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
; STREET: 3401 Hillview Ave.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/449,447  
; FILING DATE: 18-JAN-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schmonsees, William  
; REGISTRATION NUMBER: 31,796  
; REFERENCE/DOCKET NUMBER: 27610-P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-6593  
; TELEFAX: 415-496-3529  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
US-08-449-447-21

Query Match 0.5%; Score 6; DB 2; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 RSLSS 107  
Db 21 RSLSS 26

## RESULT 144

US-08-184-328-21  
; Sequence 21, Application US/08184328  
; Patent No. 5874086  
; GENERAL INFORMATION:  
; APPLICANT: Kristenansky, John L.  
; APPLICANT: Nestor Jr., John J.  
; APPLICANT: Ho, Teresa H.  
; APPLICANT: Vickery, Brian H.  
; APPLICANT: Bach, Chinh T.  
; TITLE OF INVENTION: ANALOGS OF PARATHYROID HORMONE AND  
; PARATHYROID HORMONE RELATED PEPTIDE: SYNTHESIS AND USE  
; FOR THE TREATMENT OF OSTEOPOROSIS  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
; STREET: 3401 Hillview Ave.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/184,328  
; FILING DATE: 18-JAN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 27610-P1  
TELEPHONE: 415-855-6593  
TELEFAX: 415-496-3529  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
US-08-184-328-21

Query Match 0.5%; Score 6; DB 2; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 RSLSS 107  
Db 21 RSLSS 26

RESULT 145  
US-08-521-097-21  
Sequence 21, Application US/08521097  
Patent No. 5977070  
GENERAL INFORMATION:  
APPLICANT: Kristenansky, John L.  
APPLICANT: Nestor Jr., John J.  
APPLICANT: Ho, Teresa H.  
APPLICANT: Vickery, Brian H.  
APPLICANT: Bach, Chinh T.  
TITLE OF INVENTION: ANALOGS OF PARATHYROID HORMONE AND  
TITLE OF INVENTION: PARATHYROID HORMONE RELATED PEPTIDE: SYNTHESIS AND USE  
FOR THE TREATMENT OF OSTEOPOROSIS  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patent Dept., Syntex (U.S.A.), Inc.  
STREET: 3401 Hillview Ave.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/521,097  
FILING DATE: 29-AUG-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/184,328  
FILING DATE: 18-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 27610-P1  
TELEPHONE: 415-855-6593  
TELEFAX: 415-496-3529  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO

FRAGMENT TYPE: N-terminal  
US-08-521-097-21  
Query Match 0.5%; Score 6; DB 2; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 102 RSLSS 107  
Db 21 RSLSS 26  
RESULT 146  
US-08-290-448A-30  
Sequence 30, Application US/08290448A  
Patent No. 5676954  
GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Rafnar, Thorunn  
APPLICANT: Kuo, Mei-Chang  
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,448A  
FILING DATE: August 15, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018CN  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-290-448A-30

Query Match 0.5%; Score 6; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
Db 25 MNAGVL 30

RESULT 147  
US-08-290-448A-30  
Sequence 30, Application US/08290448A

Patent No. 5698204  
; GENERAL INFORMATION:  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Klapper, David G.  
; APPLICANT: Rafnar, Thorunn  
; APPLICANT: Kuo, Mei-chang  
; TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US 07/529,951  
; FILING DATE: May 29, 1990  
; PRIORITY DATE: August 15, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/529,951  
; FILING DATE: May 29, 1990  
; PRIORITY DATE: August 15, 1994  
; NAME: Amy E. Mandragouras  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IMI-018CN  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-290-448A-30

Query Match 0.5%; Score 6; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
DB 25 MNAGVL 30

RESULT 148  
US-08-175-069A-30  
; Sequence 30, Application US/08175069A  
; Patent No. 5776761  
; GENERAL INFORMATION:  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Klapper, David G.  
; APPLICANT: Rafnar, Thorunn  
; APPLICANT: Kuo, Mei-chang  
; TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,069A  
FILING DATE: December 29, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018DV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-175-069A-30

Query Match 0.5%; Score 6; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
DB 25 MNAGVL 30

RESULT 149  
US-08-290-448A-31  
; Sequence 31, Application US/08290448A  
; Patent No. 5676954  
; GENERAL INFORMATION:  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Klapper, David G.  
; APPLICANT: Rafnar, Thorunn  
; APPLICANT: Kuo, Mei-chang  
; TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US 08/290,448A  
; FILING DATE: August 15, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/529,951  
; FILING DATE: May 29, 1990  
; APPLICATION NUMBER: US 07/325,365  
; FILING DATE: March 17, 1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Amy E. Mandragouras  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IMI-018CN

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-290-448A-31

Query Match 0.5%; Score 6; DB 1; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
DB 25 MNAGVL 30

RESULT 150  
US-08-290-448A-31  
Sequence 31, Application US/08290448A  
Patent No. 5698204  
GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Rafnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290.448A  
FILING DATE: August 15, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-290-448A-31

Query Match 0.5%; Score 6; DB 1; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
DB 25 MNAGVL 30

RESULT 151  
US-08-175-069A-31  
Sequence 31, Application US/08175069A  
Patent No. 5776761  
GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Rafnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 60 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/175.069A  
FILING DATE: December 29, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018DV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-175-069A-31

Query Match 0.5%; Score 6; DB 1; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
DB 25 MNAGVL 30

RESULT 152  
US-08-190-802A-169  
Sequence 169, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: IEF SSP 9306 rIII, Fig. 29
US-08-190-802A-169

Query Match 0.5%; Score 6; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 VDAHTA 512
Db 4 VDAHTA 9

RESULT 153
US-08-190-802A-214
Sequence 214, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.

```

```

REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: RETINOBLASTOMA BINDING PROTEIN -
HUMAN rIII, Fig. 41
US-08-190-802A-214

Query Match 0.5%; Score 6; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 VDAHTA 512
Db 4 VDAHTA 9

RESULT 154
US-08-290-448A-29
Sequence 29, Application US/08290448A
Patent No. 5676954
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,448A
FILING DATE: August 15, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal

```

US-08-290-448A-29

Query Match 0.5%; Score 6; DB 1; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
| | | | |  
Db 25 MNAGVL 30

RESULT 155

US-08-290-448A-29  
; Sequence 29, Application US/08290448A  
; Patent No. 5698204  
; GENERAL INFORMATION:  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Klapper, David G.  
; APPLICANT: Rafnar, Thorunn  
; APPLICANT: Kuo, Mei-Chang  
; TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/290,448A  
; FILING DATE: August 15, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/529,951  
; FILING DATE: May 29, 1990  
; APPLICATION NUMBER: US 07/325,365  
; FILING DATE: March 17, 1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Amy E. Mandragouras  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IMI-018CN  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: Internal  
US-08-290-448A-29

Query Match 0.5%; Score 6; DB 1; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
| | | | |  
Db 25 MNAGVL 30

RESULT 156

US-08-175-069A-29  
; Sequence 29, Application US/08175069A  
; Patent No. 5776761

; GENERAL INFORMATION:

; APPLICANT: Rogers, Bruce  
; APPLICANT: Klapper, David G.  
; APPLICANT: Rafnar, Thorunn  
; APPLICANT: Kuo, Mei-Chang  
; TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/175,069A  
; FILING DATE: December 29, 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/529,951  
; FILING DATE: May 29, 1990  
; APPLICATION NUMBER: US 07/325,365  
; FILING DATE: March 17, 1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Amy E. Mandragouras  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IMI-018DV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: Internal  
US-08-175-069A-29

Query Match 0.5%; Score 6; DB 1; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1218 MNAGVL 1223  
| | | | |  
Db 25 MNAGVL 30

RESULT 157

US-09-007-905-48  
; Sequence 48, Application US/09007905  
; Patent No. 6221585  
; GENERAL INFORMATION:  
; APPLICANT: Fran ois J.-M. Iris and Jean-Louis Pourny  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES UNDERLYING  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS

```
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007.905
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Antler, Adriane M.
; REGISTRATION NUMBER: 32,605
; REFERENCE/DOCKET NUMBER: 9408-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-007-905-48

Query Match 0.5%; Score 6; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1165 SVGSY 1170
Db 22 SVGSY 27

RESULT 158.
US-08-531-927B-5
; Sequence 5, Application US/08531927B
; Patent No. 5840491
; GENERAL INFORMATION:
; APPLICANT: Kakizuka, Akira
; TITLE OF INVENTION: DNA Sequence Encoding the Machado-Joseph
; TITLE OF INVENTION: Disease Gene and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531.927B
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP H6-251600
; FILING DATE: 21-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: ATH95-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-531-927B-5

Query Match 0.5%; Score 6; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1113 GALGSD 1118
Db 19 GALGSD 24

RESULT 159
US-09-007-905-57
; Sequence 57, Application US/09007905
; Patent No. 6221585
; GENERAL INFORMATION:
; APPLICANT: Fran ois J.-M. Iris and Jean-Louis Pourny
; TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES UNDERLYING
; TITLE OF INVENTION: DEFINED PHENOTYPES
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007.905
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Antler, Adriane M.
; REGISTRATION NUMBER: 32,605
; REFERENCE/DOCKET NUMBER: 9408-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-007-905-57
```

```
Query Match 0.5%; Score 6; DB 4; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1165 SVGSY 1170
Db 28 SVGSY 33

RESULT 160
US-08-637-759B-358
; Sequence 358, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
```

```

; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 358:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 56 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-637-759B-358

Query Match 0.5%; Score 6; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 DLYRSL 104
Db 24 DLYRSL 29

RESULT 161
US-08-871-355A-358
; Sequence 358, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 358:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 56 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-871-355A-358

Query Match 0.5%; Score 6; DB 3; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 DLYRSL 104
Db 24 DLYRSL 29

RESULT 162
US-08-468-011A-9
; Sequence 9, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

MOLECULE TYPE: protein  
US-08-468-011A-9

Query Match 0.5%; Score 6; DB 3; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 YLHNL1 1280  
|||||  
Db 47 YLHNL1 52

## RESULT 163

US-08-905-223-388  
; Sequence 388, Application US/08905223  
; Patent No. 622029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Lacroix, Bruno  
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
; NUMBER OF SEQUENCES: 503  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/905,223  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 388:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 66 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORGANISM: Homo Sapiens  
; TISSUE TYPE: Brain  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: -57..-1  
; IDENTIFICATION METHOD: Von Heijne matrix  
; OTHER INFORMATION: score 3.8  
; OTHER INFORMATION: seq QLXLVMEFCGAGS/VT

Query Match 0.5%; Score 6; DB 4; Length 66;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 DLKNT 732  
|||||  
Db 60 DLKNT 65

## RESULT 164

US-08-330-638D-2  
; Sequence 2, Application US/08330638D  
; Patent No. 5731425  
; GENERAL INFORMATION:  
; APPLICANT: Brizzard, Billy  
; APPLICANT: Bianca, Darlene  
; APPLICANT: Chubet, Richard  
; APPLICANT: Vizard, Douglas  
; APPLICANT: Hopp, Thomas  
; TITLE OF INVENTION: POLYPEPTIDE SURFACE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eastman Kodak Company,  
; STREET: 343 State Street  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14650-2201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch,  
; MEDIUM TYPE: 1.44 MB storage, (Hewlett Packard)  
; COMPUTER: HP Vectra  
; OPERATING SYSTEM: MS-DOS Version 6.0  
; SOFTWARE: WORD FOR WINDOWS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/330,638D  
; FILING DATE: 28 OCT 1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: NONE  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kiernan, Anne B.  
; REGISTRATION NUMBER: 36,566  
; REFERENCE/DOCKET NUMBER: 71255  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 588-2405  
; TELEFAX: (716) 477-4646  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 68 AMINO ACID RESIDUES  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE: SYNTHETICALLY PREPARED  
; IMMEDIATE SOURCE: SYNTHETICALLY PREPARED  
; FEATURE: SECRETION SEQUENCE  
; FEATURE: LOCATION: 1-15  
; FEATURE: FEATURE: CELL MARKER SEGMENT  
; FEATURE: LOCATION: 16-23  
; FEATURE: FEATURE: SPACER SEGMENT  
; FEATURE: LOCATION: 24-44  
; FEATURE: FEATURE: TRANSMEMBRANE SEGMENT  
; FEATURE: LOCATION: 45-65  
; FEATURE: FEATURE: ANCHOR SEGMENT  
; FEATURE: LOCATION: 66-68  
; PUBLICATION INFORMATION: NONE  
US-08-330-638D-2

Query Match 0.5%; Score 6; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGA 22  
|||||  
Db 7 LALVGA 12

## RESULT 165

US-08-906-746A-2  
; Sequence 2, Application US/08906746A  
; Patent No. 5945292

## GENERAL INFORMATION:

APPLICANT: Brizzard, Billy L.  
APPLICANT: Bianca, Darlene W.  
APPLICANT: Chubert, Richard G.  
APPLICANT: Vizard, Douglas L.  
APPLICANT: Hopp, Thomas P.  
TITLE OF INVENTION: Method of Identifying Cells with  
; TITLE OF INVENTION: Polypeptide Surface Marker  
NUMBER OF SEQUENCES: 10

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Senniger, Powers, Leavitt & Roedel  
STREET: One Metropolitan Square- 16th Floor  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63102

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/906,746A  
FILING DATE: 06-AUG-1997  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Stone, Paul A.  
REGISTRATION NUMBER: 38,628  
REFERENCE/DOCKET NUMBER: SGM 6874  
TELEPHONE: 314-231-5400  
TELEFAX: 314-231-4342

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:

NAME/KEY: Domain  
LOCATION: 1..15  
OTHER INFORMATION: /note= "Secretion Sequence"

FEATURE:  
NAME/KEY: Domain  
LOCATION: 16..23  
OTHER INFORMATION: /note= "Cell Marker Segment"

FEATURE:  
NAME/KEY: Domain  
LOCATION: 24..44  
OTHER INFORMATION: /note= "Spacer Segment"

FEATURE:  
NAME/KEY: Domain  
LOCATION: 45..65  
OTHER INFORMATION: /note= "Transmembrane Segment"

FEATURE:  
NAME/KEY: Domain  
LOCATION: 66..68  
OTHER INFORMATION: /note= "Anchor Segment"

US-08-906-746A-2  
Query Match 0.5%; Score 6; DB 2; Length 68;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGA 22

Db 7 LALVGA 12

## RESULT 166

US-08-905-223-365  
; Sequence 365, Application US/08905223  
; Patent No. 6222029  
GENERAL INFORMATION:  
APPLICANT: Edwards, Jean-Baptiste D.  
APPLICANT: Dueiert, Aymeric  
APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905,223  
FILING DATE:  
CLASSIFICATION: 536

## ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 365:

## SEQUENCE CHARACTERISTICS:

LENGTH: 73 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Brain  
FEATURE:

NAME/KEY: sig\_peptide  
LOCATION: -33..-1  
IDENTIFICATION METHOD: Von Heijne matrix  
OTHER INFORMATION: score 4.3  
OTHER INFORMATION: seq LLGWGLNLTLOG/AP  
US-08-905-223-365

Query Match 0.5%; Score 6; DB 4; Length 73;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 653 NLTGQ 658

Db 27 NLTGQ 32

## RESULT 167

US-08-422-613-4  
; Sequence 4, Application US/08422613  
; Patent No. 5695389

## GENERAL INFORMATION:

APPLICANT: Louvard, Daniel  
APPLICANT: Dudouet, Brigitte  
APPLICANT: Robine, Sylvie  
APPLICANT: Arpin, Monique

APPLICANT: Pringault, Eric  
APPLICANT: Garcia, Alphonse  
TITLE OF INVENTION: Agents for the In Vitro Diagnosis of  
Malignant Cells Originating in the Digestive Tract  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,613  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/989,696  
FILING DATE: 14-DEC-1992  
APPLICATION NUMBER: US 07/662,992  
FILING DATE: 28-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/FR86.00150  
FILING DATE: 30-APR-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 85.06707  
FILING DATE: 02-MAY-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 85.16820  
FILING DATE: 13-NOV-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 02356-0006-04000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 75 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-422-613-4

Query Match 0.5%; Score 6; DB 1; Length 75;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 948 TATTTL 953  
|||||  
Db 3 TATTTL 8

RESULT 168  
US-08-353-476-79  
Sequence 79, Application US/08353476  
Patent No. 5871902  
GENERAL INFORMATION:  
APPLICANT: Weininger, Susan  
APPLICANT: Weininger, Arthur M  
TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A  
SPECIFIC SEQUENCE COMPOSITION  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik & Saliwanchik

STREET: 2421 N.W. 41st St., Suite A-1  
CITY: Gainesville  
STATE: Florida  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/353,476  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bencen, Gerard H  
REGISTRATION NUMBER: 35,746  
REFERENCE/DOCKET NUMBER: GP-100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (904) 375-8100  
TELEFAX: (904) 372-5800  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
US-08-353-476-79

Query Match 0.5%; Score 6; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 LGQFNG 172  
|||||  
Db 78 LGQFNG 83

RESULT 169  
US-07-691-191A-6  
Sequence 6, Application US/07691191A  
Patent No. 5328990  
GENERAL INFORMATION:  
APPLICANT: Wistow, Graeme J.  
TITLE OF INVENTION: ISOLATION OF MACROPHAGE MIGRATION  
INHIBITION FACTOR FROM OCULAR LENS AND DNA WHICH ENCODES  
THE FACTOR  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 301 No. 5328990th Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22046-3487  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/691,191A  
FILING DATE: 19910426  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1173-301p  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 241-1300  
TELEFAX: (703) 241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 97 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-691-191A-6

Query Match 0.5%; Score 6; DB 1; Length 97;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 126 GQOQNK 131  
|||||  
Db 59 GQOQNK 64

## RESULT 170

US-08-202-486-6  
Sequence 6, Application US/08202486  
Patent No. 5656737

## GENERAL INFORMATION:

APPLICANT: Wistow, Graeme J.

TITLE OF INVENTION: ISOLATION OF MACROPHAGE MIGRATION

TITLE OF INVENTION: INHIBITION FACTOR FROM OCULAR LENS AND DNA WHICH ENCODES

TITLE OF INVENTION: THE FACTOR

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: 301 No. 5656737th Washington Street

CITY: Falls Church

STATE: Virginia

COUNTRY: United States

ZIP: 22046-3487

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/202.486

FILING DATE: 28-FEB-1994

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/691,191

FILING DATE: 26-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 1173-301P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 241-1300

TELEFAX: (703) 241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 97 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-202-486-6

Query Match 0.5%; Score 6; DB 1; Length 97;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 126 GQOQNK 131  
|||||

Db 59 GQOQNK 64

## RESULT 171

US-09-276-071-2

Sequence 2, Application US/09276071

Patent No. 6207149

GENERAL INFORMATION:

APPLICANT: Tsuchiya, Rie

APPLICANT: Fugisang, Claus Crone

TITLE OF INVENTION: Starch Binding Domains (SBDs) For Oral Care Products

FILE REFERENCE: 5017.204-US

CURRENT APPLICATION NUMBER: US/09/276.071

CURRENT FILING DATE: 1999-03-25

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 109

TYPE: PRT

ORGANISM: Steatothermophilus

US-09-276-071-2

Query Match 0.5%; Score 6; DB 4; Length 109;  
Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 451 TGNITV 456  
|||||

Db 100 TGNITV 105

## RESULT 172

US-07-691-191A-9

Sequence 9, Application US/07691191A

Patent No. 5328990

GENERAL INFORMATION:

APPLICANT: Wistow, Graeme J.

TITLE OF INVENTION: ISOLATION OF MACROPHAGE MIGRATION

TITLE OF INVENTION: INHIBITION FACTOR FROM OCULAR LENS AND DNA WHICH ENCODES

TITLE OF INVENTION: THE FACTOR

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: 301 No. 5328990th Washington Street

CITY: Falls Church

STATE: Virginia

COUNTRY: United States

ZIP: 22046-3487

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/691.191A

FILING DATE: 19910426

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 1173-301P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 241-1300

TELEFAX: (703) 241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 114 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein



US-07-691-191A-9

Query Match 0.5%; Score 6; DB 1; Length 114;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 126 GGOQNK 131  
Db 68 GGOQNK 73

RESULT 173

US-08-202-486-9  
; Sequence 9, Application US/08202486  
; Patent No. 5656737  
; GENERAL INFORMATION:  
; APPLICANT: Wistow, Graeme J.  
; TITLE OF INVENTION: ISOLATION OF MACROPHAGE MIGRATION  
; TITLE OF INVENTION: INHIBITION FACTOR FROM OCULAR LENS AND DNA WHICH ENCODES  
; TITLE OF INVENTION: THE FACTOR  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: 301 No. 5656737th Washington Street  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22046-3487  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/202,486  
; FILING DATE: 28-FEB-1994  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/691,191  
; FILING DATE: 26-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murphy Jr., Gerald M.  
; REGISTRATION NUMBER: 28,977  
; REFERENCE/DOCKET NUMBER: 1173-301P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 241-1300  
; TELEFAX: (703) 241-2848  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 114 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: Single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-202-486-9

Query Match 0.5%; Score 6; DB 1; Length 114;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 126 GGOQNK 131  
Db 68 GGOQNK 73

RESULT 174

US-08-558-735-8  
; Sequence 8, Application US/08558735  
; Patent No. 5681724  
; GENERAL INFORMATION:

; APPLICANT: Tripp, Cynthia A.  
; APPLICANT: Brandt, Kevin S.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: PARASITIC HELMINTH MACROPHAGE MIGRATION  
; TITLE OF INVENTION: INHIBITORY FACTOR PROTEINS, NUCLEIC ACID MOLECULES, AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/558,735  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-43  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 114 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-558-735-8

Query Match 0.5%; Score 6; DB 1; Length 114;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 KASSTV 219  
Db 21 KASSTV 26

RESULT 175

US-08-906-480-8  
; Sequence 8, Application US/08906480  
; Patent No. 6207158  
; GENERAL INFORMATION:  
; APPLICANT: Tripp, Cynthia A.  
; APPLICANT: Brandt, Kevin S.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: PARASITIC HELMINTH MACROPHAGE MIGRATION  
; TITLE OF INVENTION: INHIBITORY FACTOR PROTEINS, NUCLEIC ACID MOLECULES, AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/906.480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION NUMBER: 08/558.735  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-43  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 114 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-906-480-8

Query Match 0.5%; Score 6; DB 4; Length 114;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 KASSTV 219  
Db 21 KASSTV 26

RESULT 176  
US-07-942-245-22  
Sequence 22, Application US/07942245  
Patent No. 5639641  
GENERAL INFORMATION:  
APPLICANT: PEDERSEN, Jan T.  
APPLICANT: SEARLE, Stephen M.J.  
APPLICANT: REES, Anthony R.  
APPLICANT: ROGUSKA, Michael A.  
APPLICANT: GUILD, Braydon C.  
TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT  
TITLE OF INVENTION: ANTIBODIES  
NUMBER OF SEQUENCES: 522  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sughrue, Micon, Zinn, Macpeak & Seas  
STREET: 2100 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: United States  
ZIP: 20037-3202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: HP 9000/700 Workstation  
OPERATING SYSTEM: UNIX  
SOFTWARE: In house  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/942,245  
FILING DATE: 09-SEP-1992  
CLASSIFICATION: 530  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 293-7060  
TELEFAX: (202) 293-7860  
TELEX: 6491103  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-942-245-22

Query Match 0.5%; Score 6; DB 1; Length 115;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 ASSSVK 252  
Db 14 ASSSVK 19

RESULT 177  
US-08-558-735-2  
Sequence 2, Application US/08558735  
Patent No. 5681724  
GENERAL INFORMATION:  
APPLICANT: Tripp, Cynthia A.  
APPLICANT: Brandt, Kevin S.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: PARASITIC HELMINTH MACROPHAGE MIGRATION  
TITLE OF INVENTION: INHIBITORY FACTOR PROTEINS, NUCLEIC ACID MOLECULES, AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/558,735  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-43  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-558-735-2

Query Match 0.5%; Score 6; DB 1; Length 115;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 KASSTV 219  
Db 22 KASSTV 27

RESULT 178  
US-08-558-735-5  
Sequence 5, Application US/08558735  
Patent No. 5681724  
GENERAL INFORMATION:  
APPLICANT: Tripp, Cynthia A.  
APPLICANT: Brandt, Kevin S.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: PARASITIC HELMINTH MACROPHAGE MIGRATION  
TITLE OF INVENTION: INHIBITORY FACTOR PROTEINS, NUCLEIC ACID MOLECULES, AND  
TITLE OF INVENTION: USES THEREOF

NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/558,735  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-43  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-558-735-5

Query Match 0.5%; Score 6; DB 1; Length 115;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 KASSTV 219  
DB 22 KASSTV 27

RESULT 179  
US-08-906-480-2  
Sequence 2, Application US/08906480  
Patent No. 6207158  
GENERAL INFORMATION:  
APPLICANT: Tripp, Cynthia A.  
APPLICANT: Brandt, Kevin S.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: PARASITIC HELMINTH MACROPHAGE MIGRATION  
TITLE OF INVENTION: INHIBITORY FACTOR PROTEINS, NUCLEIC ACID MOLECULES, AND  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/906,480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/558,735  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-43  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-906-480-2

Query Match 0.5%; Score 6; DB 4; Length 115;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 KASSTV 219  
DB 22 KASSTV 27

RESULT 180  
US-08-906-480-5  
Sequence 5, Application US/08906480  
Patent No. 6207158  
GENERAL INFORMATION:  
APPLICANT: Tripp, Cynthia A.  
APPLICANT: Brandt, Kevin S.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: PARASITIC HELMINTH MACROPHAGE MIGRATION  
TITLE OF INVENTION: INHIBITORY FACTOR PROTEINS, NUCLEIC ACID MOLECULES, AND  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/906,480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/558,735  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-43  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-906-480-5

Query Match 0.5%; Score 6; DB 4; Length 115;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 KASSTV 219  
111111  
Db 22 KASSTV 27

## RESULT 181

US-08-800-198-2  
; Sequence 2, Application US/08800198  
; Patent No. 5942602  
; GENERAL INFORMATION:  
; APPLICANT: WELS, WINFRIED S.  
; APPLICANT: SCHMIDT, MATHIAS  
; APPLICANT: VAKALOPOULOU, EVANGELIA  
; APPLICANT: SCHNEIDER, DOUGLAS  
; TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
; STREET: 2200 CLARENDON BLVD. SUITE 1400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: US  
; ZIP: 22201

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 13-FEB-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HAMLET-KING, DIANA  
; REGISTRATION NUMBER: 33,302  
; REFERENCE/DOCKET NUMBER: SCH 1576  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-243-6333  
; TELEFAX: 703-243-6410  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 119 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-800-198-2

Query Match 0.5%; Score 6; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 YGYDFA 1154  
111111  
Db 102 YGYDFA 107

## RESULT 182

US-09-296-595-2  
; Sequence 2, Application US/09296595A  
; Patent No. 6129915  
; GENERAL INFORMATION:  
; APPLICANT: WELS, WINFRIED S.  
; APPLICANT: SCHMIDT, MATHIAS  
; APPLICANT: VAKALOPOULOU, EVANGELIA  
; APPLICANT: SCHNEIDER, DOUGLAS

; TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES  
; FILE REFERENCE: SCH-1576 D1  
; CURRENT APPLICATION NUMBER: US/09/296,595A  
; CURRENT FILING DATE: 1999-04-23  
; EARLIER APPLICATION NUMBER: 08/800,198  
; EARLIER FILING DATE: 1997-02-13  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Murine sp.  
US-09-296-595-2

Query Match 0.5%; Score 6; DB 3; Length 119;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 YGYDFA 1154  
111111  
Db 102 YGYDFA 107

## RESULT 183

US-08-844-188-36  
; Sequence 36, Application US/08844188  
; Patent No. 6127180  
; GENERAL INFORMATION:  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Knuth, Mark  
; APPLICANT: Pollard, Michael R.  
; APPLICANT: Cardineau, Guy  
; APPLICANT: Schwab, George E.  
; TITLE OF INVENTION: Pesticidal Toxins  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606-6669  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/844,188  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/633,993  
; FILING DATE: 19-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sanders, Jay M.  
; REGISTRATION NUMBER: 39,355  
; REFERENCE/DOCKET NUMBER: MA-703C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-375-8100  
; TELEFAX: 352-372-5800  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-844-188-36

Query Match 0.5%; Score 6; DB 3; Length 123;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AEISLY 239  
| | | | |  
DB 69 AEISLY 74

## RESULT 184

US-08-844-188-41  
; Sequence 41, Application US/08844188  
; Patent No. 6127180

## GENERAL INFORMATION:

; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Knuth, Mark  
; APPLICANT: Pollard, Michael R.  
; APPLICANT: Cardineau, Guy  
; APPLICANT: Schwab, George E.  
; TITLE OF INVENTION: Pesticidal Toxins  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606-6669

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/844,188  
; FILING DATE:  
; CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/633,993  
; FILING DATE: 19-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sanders, Jay M.  
; REGISTRATION NUMBER: 39,355  
; REFERENCE/DOCKET NUMBER: MA-703C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-375-8100  
; TELEFAX: 352-372-5800

## INFORMATION FOR SEQ ID NO: 41:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-844-188-41

Query Match 0.5%; Score 6; DB 3; Length 123;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AEISLY 239  
| | | | |  
DB 69 AEISLY 74

## RESULT 185

US-09-156-316-7  
; Sequence 7, Application US/09156316  
; Patent No. 6183961

## GENERAL INFORMATION:

; APPLICANT: Bernstein, Harold S.  
; APPLICANT: Coughlin, Shaun R.  
; TITLE OF INVENTION: Methods and Compositions for Regulating Cell Cycle  
; FILE REFERENCE: UCSF-020/0105  
; CURRENT APPLICATION NUMBER: US/09/156,316  
; CURRENT FILING DATE: 1998-09-18  
; EARLIER FILING DATE: 1997-09-22  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 123  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-156-316-7

Query Match 0.5%; Score 6; DB 4; Length 123;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 DFGSVT 533  
| | | | |  
DB 14 DFGSVT 19

## RESULT 186

US-08-276-852-54  
; Sequence 54, Application US/08276852  
; Patent No. 5652138

## GENERAL INFORMATION:

; APPLICANT: Burton, Dennis R  
; APPLICANT: Barbas, Carlos F  
; APPLICANT: Lerner, Richard A  
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES  
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS  
; NUMBER OF SEQUENCES: 170  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,  
; STREET: Mail Drop TPC8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/276,852  
; FILING DATE: 18-JUL-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/178,302  
; FILING DATE: 30-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/954,148  
; FILING DATE: 30-SEP-1992  
; NAME: Fitting, Thomas  
; REGISTRATION NUMBER: 34,163  
; REFERENCE/DOCKET NUMBER: SCR1452P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-554-2937  
; TELEFAX: 619-554-6312  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 124 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-276-852-54

Query Match 0.5%; Score 6; DB 1; Length 124;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGAL 23  
Db 110 ALVGAL 115

## RESULT 187

US-08-899-575-54  
Sequence 54, Application US/08899575  
Patent No. 5770440  
GENERAL INFORMATION:  
APPLICANT: Burton, Dennis R  
APPLICANT: Barbas, Carlos F  
APPLICANT: Lerner, Richard A  
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES  
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS  
NUMBER OF SEQUENCES: 170  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESS: Patent Counsel  
STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,  
STREET: Mail Drop TPC8  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,575  
FILING DATE: 24-JUL-1997  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/276,852  
FILING DATE: 18-JUL-1994  
APPLICATION NUMBER: US 08/178,302  
FILING DATE: 30-SEP-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/954,148  
FILING DATE: 30-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: SCRI452P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 124 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-899-575-54

Query Match 0.5%; Score 6; DB 1; Length 124;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGAL 23

Db 110 ALVGAL 115

## RESULT 188

US-08-899-575-54  
Sequence 54, Application US/08899575  
Patent No. 5804440  
GENERAL INFORMATION:  
APPLICANT: Burton, Dennis R  
APPLICANT: Barbas, Carlos F  
APPLICANT: Lerner, Richard A  
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES  
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS  
NUMBER OF SEQUENCES: 170  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESS: Patent Counsel  
STREET: 10666 No. 5804440th Torrey Pines Road, Suite 220,  
STREET: Mail Drop TPC8  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,575  
FILING DATE: 24-JUL-1997  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/276,852  
FILING DATE: 18-JUL-1994  
APPLICATION NUMBER: US 08/178,302  
FILING DATE: 30-SEP-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/954,148  
FILING DATE: 30-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: SCRI452P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 124 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-899-575-54

Query Match 0.5%; Score 6; DB 1; Length 124;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGAL 23  
Db 110 ALVGAL 115

## RESULT 189

US-08-630-172-4  
Sequence 4, Application US/08630172  
Patent No. 6060054  
GENERAL INFORMATION:  
APPLICANT: Staerz, Uwe  
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T

;; TITLE OF INVENTION: LYMPHOCYTE VETO  
;; NUMBER OF SEQUENCES: 41  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sheridan Ross & McIntosh  
;; STREET: 1700 Lincoln Street, 35th Floor  
;; CITY: Denver  
;; STATE: Colorado  
;; COUNTRY: U.S.  
;; ZIP: 80203

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/630,172  
;; FILING DATE:

;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:

;; NAME: Connell, Gary J.  
;; REGISTRATION NUMBER: 32,020  
;; REFERENCE/DOCKET NUMBER: 2879-36  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (303) 863-9700  
;; TELEFAX: (303) 863-0223

;; INFORMATION FOR SEQ ID NO: 4:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 124 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; US-08-630-172-4

Query Match 0.5%; Score 6; DB 3; Length 124;  
Best Local Similarity 100.0%; Pred. No. 6e+02; Indels 0; Mismatches 0; Gaps 0;

QY 669 NLTIQ 674

Db 76 NLTIQ 81

RESULT 190

PCT-US95-08743-54

;; Sequence 54, Application PC/TUS9508743

;; GENERAL INFORMATION:

;; APPLICANT:

;; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES

;; NUMBER OF SEQUENCES: 170

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: PCT/US95/08743

;; FILING DATE: 11-JUL-1995

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/276,852

;; FILING DATE: 18-JUL-1994

;; INFORMATION FOR SEQ ID NO: 54:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 124 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; PCT-US95-08743-54

Query Match 0.5%; Score 6; DB 5; Length 124;  
Best Local Similarity 100.0%; Pred. No. 6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGL 23

Db 110 ALVGL 115

RESULT 191

US-08-284-516C-57

;; Sequence 57, Application US/08284516C

;; Patent No. 6056957

;; GENERAL INFORMATION:

;; APPLICANT: Chou, Chuan-Chu

;; APPLICANT: Murgolo, Nicholas

;; APPLICANT: Abrams, John

;; APPLICANT: Jenh, Chung-Her

;; APPLICANT: Petro, Mary

;; APPLICANT: Silver, Jon

;; APPLICANT: Tindall, Stephen

;; APPLICANT: Windsor, William

;; APPLICANT: Zavodny, Paul

;; TITLE OF INVENTION: Design, Cloning and Expression of Humanized Monoclonal Ant

;; NUMBER OF SEQUENCES: 66

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Schering-Plough Corporation

;; STREET: 2000 Galloping Hill Road

;; CITY: Kenilworth

;; STATE: New Jersey

;; COUNTRY: USA

;; ZIP: 07033-0530

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: Apple Macintosh

;; OPERATING SYSTEM: Macintosh 6.0.5

;; SOFTWARE: Microsoft Word 4.00B

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/284,516C

;; FILING DATE: 04-AUG-1994

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/US93/00759

;; FILING DATE: 04-FEB-1993

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: 07/832,842

;; FILING DATE: 06-FEB-1992

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Foulke, Cynthia L.

;; REGISTRATION NUMBER: 32,364

;; REFERENCE/DOCKET NUMBER: JB0233K

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 908 298 2987

;; TELEFAX: 908 298 5388

;; INFORMATION FOR SEQ ID NO: 57:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 127 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; US-08-284-516C-57

Query Match 0.5%; Score 6; DB 3; Length 127;

Best Local Similarity 100.0%; Pred. No. 6.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 QASEGI 228

Db 43 QASEGI 48

RESULT 192

US-08-276-852-57

;; Sequence 57, Application US/08276852

;; Patent No. 5652138

;; GENERAL INFORMATION:  
;; APPLICANT: Burton, Dennis R  
;; APPLICANT: Barbas, Carlos F  
;; APPLICANT: Lerner, Richard A  
;; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES  
;; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS  
;; NUMBER OF SEQUENCES: 170  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: The Scripps Research Institute, Office of  
;; ADDRESSEE: Patent Counsel  
;; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,  
;; STREET: Mail Drop TPC8  
;; CITY: La Jolla  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 92037  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0; Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/276,852  
;; FILING DATE: 18-JUL-1994  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/178,302  
;; FILING DATE: 30-SEP-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/954,148  
;; FILING DATE: 30-SEP-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitting, Thomas  
;; REGISTRATION NUMBER: 34,163  
;; REFERENCE/DOCKET NUMBER: SCR1452P  
;; TELEPHONE: 619-554-2937  
;; TELEFAX: 619-554-6312  
;; INFORMATION FOR SEQ ID NO: 57:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 128 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-276-852-57

Query Match 0.5%; Score 6; DB 1; Length 128;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGAL 23  
Db 110 ALVGAL 115

RESULT 193  
US-08-899-575-57  
; Sequence 57, Application US/08899575  
; Patent No. 5770440  
; GENERAL INFORMATION:  
; APPLICANT: Burton, Dennis R  
; APPLICANT: Barbas, Carlos F  
; APPLICANT: Lerner, Richard A  
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES  
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS  
; NUMBER OF SEQUENCES: 170  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,  
; STREET: Mail Drop TPC8  
; CITY: La Jolla

;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 92037  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0; Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/899,575  
;; FILING DATE: 24-JUL-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/276,852  
;; FILING DATE: 18-JUL-1994  
;; APPLICATION NUMBER: US 08/178,302  
;; FILING DATE: 30-SEP-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/954,148  
;; FILING DATE: 30-SEP-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitting, Thomas  
;; REGISTRATION NUMBER: 34,163  
;; REFERENCE/DOCKET NUMBER: SCR1452P  
;; TELEPHONE: 619-554-2937  
;; TELEFAX: 619-554-6312  
;; INFORMATION FOR SEQ ID NO: 57:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 128 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-899-575-57

Query Match 0.5%; Score 6; DB 1; Length 128;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGAL 23  
Db 110 ALVGAL 115

RESULT 194  
US-08-899-575-57  
; Sequence 57, Application US/08899575  
; Patent No. 5804440  
; GENERAL INFORMATION:  
; APPLICANT: Burton, Dennis R  
; APPLICANT: Barbas, Carlos F  
; APPLICANT: Lerner, Richard A  
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES  
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS  
; NUMBER OF SEQUENCES: 170  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 No. 5804440th Torrey Pines Road, Suite 220,  
; STREET: Mail Drop TPC8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0; Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/899,575  
; FILING DATE: 24-JUL-1997



CLASSIFICATION:  
PRIOR APPLICATION DATA: US 08/276,852  
FILING DATE: 18-JUL-1994  
APPLICATION NUMBER: US 08/178,302  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA: US 07/954,148  
FILING DATE: 30-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: SCRI452P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 128 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-899-575-57

Query Match 0.5%; Score 6; DB 1; Length 128;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGAL 23  
|||||  
DB 110 ALVGAL 115

RESULT 195  
PCT-US95-08743-57  
Sequence 57, Application PC/TUS9508743  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES  
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS  
NUMBER OF SEQUENCES: 170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/08743  
FILING DATE: 11-JUL-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/276,852  
FILING DATE: 18-JUL-1994  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 128 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-08743-57

Query Match 0.5%; Score 6; DB 5; Length 128;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGAL 23  
|||||  
DB 110 ALVGAL 115

RESULT 196  
US-08-611-107-20

Sequence 20, Application US/08611107  
Patent No. 5801233  
GENERAL INFORMATION:  
APPLICANT: Haselkorn, Robert  
APPLICANT: Gornicki, Piotr  
TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS ENCODING  
TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE AND USES  
TITLE OF INVENTION: THEREFOR  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/611,107  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US SN 07/956,700  
FILING DATE: 02-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US93/09340  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US SN 08/422,560  
FILING DATE: 14-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: ARCD:221  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-611-107-20

Query Match 0.5%; Score 6; DB 1; Length 132;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTG 452  
|||||  
DB 118 VENLTG 123

RESULT 197  
US-08-422-560A-20  
Sequence 20, Application US/08422560A  
Patent No. 5910626  
GENERAL INFORMATION:  
APPLICANT: Haselkorn, Robert  
APPLICANT: Gornicki, Piotr  
TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND  
TITLE OF INVENTION: METHODS FOR USE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston

STATE: TX  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,560A  
FILING DATE: 14-APR-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION NUMBER: US 07/956,700  
FILING DATE: 02-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilson, Mark B.  
REGISTRATION NUMBER: 37,259  
REFERENCE/DOCKET NUMBER: ARCD:152/WIM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512-418-3000  
TELEFAX: 512-474-7577  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-422-560A-20

Query Match 0.5%; Score 6; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTG 452  
|||||  
Db 118 VENLTG 123

RESULT 198  
US-08-468-793-20  
Sequence 20, Application US/08468793  
Patent No. 6177267  
GENERAL INFORMATION:  
APPLICANT: Haselkorn, Robert  
APPLICANT: Gornicki, Piotr  
TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND  
METHODS OF USE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,793  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION NUMBER: US 08/422,560  
FILING DATE: 14-APR-1995  
APPLICATION NUMBER: US SN 07/956,700  
FILING DATE: 02-OCT-1992  
CLASSIFICATION: 800  
APPLICATION NUMBER: PCT/US93/09340

FILING DATE: 30-SEP-1993  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: ARCD:152/KIT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-468-793-20

Query Match 0.5%; Score 6; DB 4; Length 132;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTG 452  
|||||  
Db 118 VENLTG 123

RESULT 199  
US-08-804-180C-2  
Sequence 2, Application US/08804180C  
Patent No. 6107056  
GENERAL INFORMATION:  
APPLICANT: Martin K. Oaks  
TITLE OF INVENTION: SCTL-4 and its Soluble Products  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Thomas M. Wozny  
STREET: 100 East Wisconsin Avenue  
CITY: Milwaukee  
STATE: Wisconsin  
COUNTRY: USA  
ZIP: 53202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.50 inch Disk  
COMPUTER: IBM  
OPERATING SYSTEM: DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,180C  
FILING DATE: February 20, 1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas M. Wozny  
REGISTRATION NUMBER: 28,922  
REFERENCE/DOCKET NUMBER: 3284-00003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (414) 271-7590  
TELEFAX: (414) 271-5770  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: no  
HYPOTHETICAL: no  
FRAGMENT TYPE: Mature Polypeptide  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapien  
DEVELOPMENTAL STAGE: Adult  
TISSUE TYPE: Lymphnode

;/ FEATURE:  
;/ NAME/KEY: Human SCTLA-4  
;/ IDENTIFICATION METHOD: Found by experiment  
;/ OTHER INFORMATION: Asn 76 and Asn 108 represent N-linked glycosylation; B7 binding  
;/ US-08-804-180C-2

Query Match 0.5%; Score 6; DB 3; Length 137;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 NLTIQG 674  
Db 76 NLTIQG 81

RESULT 200  
US-08-685-808-2  
; Sequence 2, Application US/08685808  
; Patent No. 6048715  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, CHARLES A., et al  
; TITLE OF INVENTION: SEPARATION AND CONCENTRATION SYSTEMS BASED  
; TITLE OF INVENTION: ON SOLUBLE OLIGOSACCHARIDE BINDING DOMAINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RAE-VENTER LAW GROUP  
; STREET: 260 Sheridan Ave., Ste. 440  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,808  
; FILING DATE: 24-JULY-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/505,860  
; FILING DATE: 24-JULY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rae-Venter, Barbara  
; REGISTRATION NUMBER: 32,750  
; REFERENCE/DOCKET NUMBER: CDDT-017.01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 328-4400  
; TELEFAX: (650) 328-4477  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 143 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: no

US-08-685-808-2

Query Match 0.5%; Score 6; DB 3; Length 143;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1047 SLYGTS 1052  
Db 17 SLYGTS 22

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2001, 09:34:19 ; Search time 39.88 Seconds  
(without alignments)  
4299.582 Million cell updates/sec

Title: US-09-360-934A-3  
Perfect score: 1296  
Sequence: 1 MEIQTHRKINPLVSLAV.....HNLSNIGHFASNLGMRYSF 1296

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 200 summaries

Database :

SPTREMBL\_16:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.unclassified:\*  
13: sp.vertebrate:\*  
14: sp.virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	148	11.4	148	2 Q9R399	Q9R399 helicobacte
2	143	11.0	1290	2 Q34111	Q34111 helicobacte
3	143	11.0	1295	2 Q34110	Q34110 helicobacte
4	99	7.6	1290	2 Q9ZHT1	Q9ZHT1 helicobacte
5	96	7.4	160	2 Q32663	Q32663 helicobacte
6	94	7.3	160	2 Q32660	Q32660 helicobacte
7	92	7.1	148	2 Q9R844	Q9R844 helicobacte
8	89	6.9	143	2 Q9LB57	Q9LB57 helicobacte
9	88	6.8	120	2 Q9LB53	Q9LB53 helicobacte
10	87	6.7	160	2 Q32671	Q32671 helicobacte
11	87	6.7	160	2 Q32684	Q32684 helicobacte
12	87	6.7	371	2 Q86250	Q86250 helicobacte
13	86	6.6	120	2 Q9LB51	Q9LB51 helicobacte
14	86	6.0	160	2 Q32669	Q32669 helicobacte
15	72	5.6	148	2 Q9R848	Q9R848 helicobacte
16	72	5.6	160	2 Q32661	Q32661 helicobacte
17	72	5.6	160	2 Q32674	Q32674 helicobacte
18	71	5.5	109	2 Q9LB50	Q9LB50 helicobacte
19	70	5.4	148	2 Q9R847	Q9R847 helicobacte

20	68	5.2	1290	2	Q06021	helicobacte
21	66	5.1	1323	2	Q87018	helicobacte
22	65	5.0	160	2	Q32672	helicobacte
23	65	5.0	160	2	Q32681	helicobacte
24	65	5.0	160	2	Q32687	helicobacte
25	60	4.6	96	2	Q9RE83	helicobacte
26	60	4.6	148	2	Q9R842	helicobacte
27	60	4.6	207	2	Q9KX03	helicobacte
28	60	4.6	207	2	Q9KX02	helicobacte
29	60	4.6	207	2	Q9KX01	helicobacte
30	58	4.5	96	2	Q9R305	helicobacte
31	57	4.4	91	2	Q9LBA7	helicobacte
32	57	4.4	160	2	Q32677	helicobacte
33	56	4.3	148	2	Q9R313	helicobacte
34	56	4.3	160	2	Q32683	helicobacte
35	56	4.3	862	2	Q9F9G3	helicobacte
36	55	4.2	148	2	Q9R851	helicobacte
37	55	4.0	78	2	Q9L811	helicobacte
38	52	4.0	78	2	Q9L807	helicobacte
39	52	4.0	78	2	Q9K2T9	helicobacte
40	52	4.0	96	2	Q9RE76	helicobacte
41	52	4.0	1291	2	Q9ZHU8	helicobacte
42	52	4.0	1291	2	Q9ZHU7	helicobacte
43	51	3.9	114	2	Q9LB54	helicobacte
44	51	3.9	143	2	Q9LB56	helicobacte
45	51	3.9	145	2	Q9LB48	helicobacte
46	51	3.9	148	2	Q9S316	helicobacte
47	51	3.9	148	2	Q9S313	helicobacte
48	51	3.9	148	2	Q9S311	helicobacte
49	51	3.9	148	2	Q9S3H8	helicobacte
50	51	3.9	160	2	Q32664	helicobacte
51	51	3.9	207	2	Q9KX00	helicobacte
52	50	3.9	160	2	Q32670	helicobacte
53	49	3.8	85	2	Q9LB81	helicobacte
54	48	3.7	98	2	Q9LB95	helicobacte
55	48	3.7	130	2	Q9LBAG	helicobacte
56	48	3.7	160	2	Q32676	helicobacte
57	48	3.7	1296	2	Q9R958	helicobacte
58	48	3.7	1324	2	Q9LBC7	helicobacte
59	48	3.7	1324	2	Q9LBC3	helicobacte
60	48	3.7	1324	2	Q9LBC2	helicobacte
61	47	3.6	119	2	Q9LB41	helicobacte
62	47	3.6	1291	2	Q9ZHT2	helicobacte
63	47	3.6	1296	2	Q9ZHV2	helicobacte
64	47	3.6	1296	2	Q9ZHV1	helicobacte
65	47	3.6	1328	2	Q9LBC8	helicobacte
66	46	3.5	148	2	Q9S310	helicobacte
67	46	3.5	1291	2	Q9ZHU4	helicobacte
68	46	3.5	1293	2	Q9ZHV0	helicobacte
69	46	3.5	1300	2	Q9ZHU0	helicobacte
70	45	3.5	139	2	Q9K2R6	helicobacte
71	45	3.5	160	2	Q32658	helicobacte
72	45	3.5	160	2	Q32679	helicobacte
73	45	3.5	160	2	Q32686	helicobacte
74	45	3.5	380	2	Q9R960	helicobacte
75	45	3.5	826	2	Q9F9G4	helicobacte
76	45	3.5	829	2	Q9F9G0	helicobacte
77	45	3.5	829	2	Q9F9F8	helicobacte
78	45	3.5	861	2	Q9F9G5	helicobacte
79	45	3.5	861	2	Q9F9F7	helicobacte
80	45	3.5	864	2	Q9F9G1	helicobacte
81	45	3.5	866	2	Q9F9G6	helicobacte
82	45	3.5	1082	2	Q9R964	helicobacte
83	45	3.5	1291	2	Q9ZHV3	helicobacte
84	45	3.5	1291	2	Q9ZHU9	helicobacte
85	45	3.5	1291	2	Q9ZHT9	helicobacte
86	45	3.5	1291	2	Q9ZHT5	helicobacte
87	45	3.5	1291	2	Q9R962	helicobacte
88	45	3.5	1291	2	Q9LBC1	helicobacte
89	45	3.5	1294	2	Q9R959	helicobacte
90	45	3.5	1296	2	Q9ZHT7	helicobacte
91	45	3.5	1303	2	Q9KJAG	helicobacte
92	44	3.4	148	2	Q87749	helicobacte

93	44	3.4	148	2	Q9S315	Q9S315 helicobacte	166	30	2.3	96	2	Q9RE80	Q9re80 helicobacte
94	44	3.4	148	2	Q9R853	Q9r853 helicobacte	167	30	2.3	96	2	Q9RE79	Q9re79 helicobacte
95	44	3.4	148	2	Q9R843	Q9r843 helicobacte	168	30	2.3	96	2	Q9R2Y8	Q9r2Y8 helicobacte
96	44	3.4	148	2	Q9R316	Q9r316 helicobacte	169	30	2.3	96	2	Q9R2X3	Q9r2r3 helicobacte
97	44	3.4	160	2	Q93257	Q93257 helicobacte	170	30	2.3	130	2	Q9LB49	Q9lb49 helicobacte
98	44	3.4	160	2	Q93273	Q93273 helicobacte	171	30	2.3	214	2	Q9K2M6	Q9k2m6 helicobacte
99	44	3.4	160	2	Q93275	Q93275 helicobacte	172	30	2.3	216	2	Q9L7T0	Q9l7t0 helicobacte
100	43	3.3	96	2	Q9RE74	Q9re74 helicobacte	173	30	2.3	216	2	Q9L7S8	Q9l7s8 helicobacte
101	43	3.3	96	2	Q9R3Y5	Q9r3Y5 helicobacte	174	30	2.3	216	2	Q9L7S6	Q9l7s6 helicobacte
102	43	3.3	96	2	Q9R2Q5	Q9r2Q5 helicobacte	175	30	2.3	240	2	Q9X412	Q9x412 helicobacte
103	43	3.3	130	2	Q9LB52	Q9lb52 helicobacte	176	30	2.3	244	2	Q34107	Q34107 helicobacte
104	42	3.2	78	2	Q9L808	Q9l808 helicobacte	177	30	2.3	244	2	Q34108	Q34108 helicobacte
105	42	3.2	403	2	Q86253	Q86253 helicobacte	178	30	2.3	244	2	Q9ZIE5	Q9zie5 helicobacte
106	41	3.2	148	2	Q9R850	Q9r850 helicobacte	179	29	2.2	83	2	Q9LB88	Q9lb88 helicobacte
107	41	3.2	160	2	Q93265	Q93265 helicobacte	180	29	2.2	89	2	Q9LB76	Q9lb76 helicobacte
108	41	3.2	160	2	Q93267	Q93267 helicobacte	181	29	2.2	96	2	Q9RE86	Q9re86 helicobacte
109	41	3.2	240	2	Q85022	Q85022 helicobacte	182	29	2.2	96	2	Q9RE78	Q9re78 helicobacte
110	41	3.2	240	2	Q85029	Q85029 helicobacte	183	29	2.2	96	2	Q9RE77	Q9re77 helicobacte
111	41	3.2	240	2	Q9X408	Q9x408 helicobacte	184	29	2.2	96	2	Q9R2R4	Q9r2r4 helicobacte
112	41	3.2	244	2	Q34109	Q34109 helicobacte	185	29	2.2	96	2	Q9LB60	Q9lb60 helicobacte
113	41	3.2	406	2	Q9Z487	Q9z487 helicobacte	186	29	2.2	97	2	Q9RE85	Q9re85 helicobacte
114	41	3.2	1291	2	Q87H01	Q87h01 helicobacte	187	29	2.2	98	2	Q9LB63	Q9lb63 helicobacte
115	41	3.2	1296	2	Q9ZHU3	Q9zhu3 helicobacte	188	29	2.2	118	2	Q9K331	Q9k331 helicobacte
116	40	3.1	244	2	Q87412	Q87412 helicobacte	189	29	2.2	120	2	Q9LB78	Q9lb78 helicobacte
117	40	3.1	406	2	Q87484	Q87484 helicobacte	190	29	2.2	123	2	Q9LB82	Q9lb82 helicobacte
118	40	3.1	1291	2	Q9ZHU6	Q9zhu6 helicobacte	191	29	2.2	136	2	Q9K2T4	Q9k2t4 helicobacte
119	39	3.0	160	2	Q93262	Q93262 helicobacte	192	29	2.2	139	2	Q9L8C4	Q9l8c4 helicobacte
120	39	3.0	160	2	Q93268	Q93268 helicobacte	193	29	2.2	143	2	Q9LB98	Q9lb98 helicobacte
121	38	2.9	148	2	Q9S314	Q9s314 helicobacte	194	29	2.2	149	2	Q9LB88	Q9lb88 helicobacte
122	38	2.9	148	2	Q9R849	Q9r849 helicobacte	195	29	2.2	214	2	Q9LAJ6	Q9laj6 helicobacte
123	38	2.9	160	2	Q93266	Q93266 helicobacte	196	28	2.2	148	2	Q87750	Q87750 helicobacte
124	38	2.9	160	2	Q93268	Q93268 helicobacte	197	28	2.2	148	2	Q9R845	Q9r845 helicobacte
125	38	2.9	406	2	Q87486	Q87486 helicobacte	198	28	2.2	148	2	Q9R841	Q9r841 helicobacte
126	38	2.9	406	2	Q87485	Q87485 helicobacte	199	28	2.2	148	2	Q9R840	Q9r840 helicobacte
127	38	2.9	1291	2	Q9ZHT8	Q9zht8 helicobacte	200	28	2.2	148	2	Q9R839	Q9r839 helicobacte
128	38	2.9	1291	2	Q9R961	Q9r961 helicobacte							
129	38	2.9	1298	2	Q9ZHT4	Q9zht4 helicobacte							
130	37	2.9	131	2	Q9LB84	Q9lb84 helicobacte							
131	37	2.9	148	2	Q9R846	Q9r846 helicobacte							
132	37	2.9	160	2	Q93256	Q93256 helicobacte							
133	37	2.9	160	2	Q93278	Q93278 helicobacte							
134	37	2.9	206	2	Q53040	Q53040 helicobacte							
135	36	2.8	148	2	Q9S312	Q9s312 helicobacte							
136	35	2.7	35	2	Q9R2Y5	Q9r2Y5 helicobacte							
137	35	2.7	96	2	Q9RE82	Q9re82 helicobacte							
138	35	2.7	125	2	Q9LB84	Q9lb84 helicobacte							
139	35	2.7	127	2	Q9LEC9	Q9lbc9 helicobacte							
140	35	2.7	144	2	Q9LBB3	Q9lbb3 helicobacte							
141	35	2.7	216	2	Q9L7T1	Q9l7t1 helicobacte							
142	35	2.7	216	2	Q9L7S9	Q9l7s9 helicobacte							
143	35	2.7	216	2	Q9L7S7	Q9l7s7 helicobacte							
144	35	2.7	216	2	Q9L7S5	Q9l7s5 helicobacte							
145	35	2.7	216	2	Q9L7S4	Q9l7s4 helicobacte							
146	35	2.7	223	2	Q9L7S1	Q9l7s1 helicobacte							
147	34	2.6	96	2	Q9RE81	Q9re81 helicobacte							
148	34	2.6	128	2	Q9LBC5	Q9lbc5 helicobacte							
149	33	2.5	112	2	Q9LB69	Q9lbc9 helicobacte							
150	33	2.5	127	2	Q9LBB6	Q9lbb6 helicobacte							
151	33	2.5	140	2	Q9LB79	Q9lb79 helicobacte							
152	33	2.5	144	2	Q9LB68	Q9lb68 helicobacte							
153	33	2.5	145	2	Q9LB97	Q9lb97 helicobacte							
154	33	2.5	148	2	Q9S3H9	Q9s3h9 helicobacte							
155	33	2.5	148	2	Q9S3H7	Q9s3h7 helicobacte							
156	33	2.5	148	2	Q9LBA1	Q9lba1 helicobacte							
157	33	2.5	160	2	Q93268	Q93268 helicobacte							
158	33	2.5	210	2	Q9R356	Q9r356 helicobacte							
159	32	2.5	861	2	Q9F9G2	Q9f9g2 helicobacte							
160	31	2.4	78	2	Q9L810	Q9l810 helicobacte							
161	31	2.4	118	2	Q9KI22	Q9ki22 helicobacte							
162	31	2.4	148	2	Q9R836	Q9r836 helicobacte							
163	31	2.4	214	2	Q9LAJ8	Q9laj8 helicobacte							
164	31	2.4	244	2	Q9R9B0	Q9r9b0 helicobacte							
165	30	2.3	96	2	Q9RE84	Q9re84 helicobacte							

## ALIGNMENTS

RESULT 1

Q9R3R9	PRELIMINARY;	PRT;	148 AA.
ID	Q9R3R9		
AC	Q9R3R9;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)		
DE	VACUOLATING CYTOTOXIN (FRAGMENT).		
GN	VACA.		
OS	Helicobacter pylori (Campylobacter pylori).		
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;		
OC	Helicobacter.		
ON	NCBI_TaxID=210;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-NCTC11638, AND 25;		
RX	MEDLINE-99255683; PubMed-10320570;		
RA	Achteman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,		
RA	Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;		
RT	"Recombination and clonal groupings within Helicobacter pylori from		
RT	different geographic regions."		
RL	Mol. Microbiol. 32:459-470(1999).		
DR	EMBL; AJ239563; CAB37375.1; -		
DR	EMBL; AJ239560; CAB37372.1; -		
FT	NON_TER 1		
FT	NON_TER 148		
SQ	SEQUENCE 148 AA; 16561 MW; 4B4FED850974B5AB CRC64;		

Query Match 11.4%; Score 148; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 3.1e-145;  
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EEANKTPDKPDKVRIQAGKGFNEFPNKEYDLYRSLSSKIDGGWGNAAHYWYKGGQ 128  
Db 1 EEANKTPDKPDKVRIQAGKGFNEFPNKEYDLYRSLSSKIDGGWGNAAHYWYKGGQ 60  
QY 129 QNKLEVDMKDAVGTYSGLRNFTGDLVNNQKATRLRGQFNNGSFTSYKDSADRTTRV 188  
Db 61 QNKLEVDMKDAVGTYSGLRNFTGDLVNNQKATRLRGQFNNGSFTSYKDSADRTTRV 120  
QY 189 DNANKNISIDNEVEINNRVGSAGRKAS 216  
Db 121 DNANKNISIDNEVEINNRVGSAGRKAS 148

RESULT 2  
ID O34111 PRELIMINARY; PRT; 1290 AA.  
AC O34111;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F79;  
RX MEDLINE=97339580; PubMed=9196179;  
RA Ito Y., Azuma T., Ito S., Miyaji H., Hirai M., Yamazaki Y., Sato F.,  
RA Kato T., Kohli Y., Kuriyama M.;  
RT "Analysis and typing of the vacA gene from caga-positive strains of  
RT Helicobacter pylori isolated in Japan.";  
RL J. Clin. Microbiol. 35:1710-1714(1997).  
DR EMBL; AF071097; AAC77452.1;  
SQ SEQUENCE 1290 AA; 139321 MW; CDA478D88AF30E67 CRC64;

Query Match 11.0%; Score 143; DB 2; Length 1290;  
Best Local Similarity 100.0%; Pred. No. 3.4e-139;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEIQOTHRKNRPLVSLALVGVLSITPQOASHAFTTIIIPAIIVGGIATGAVGTVSG 60  
Db 1 MEIQOTHRKNRPLVSLALVGVLSITPQOASHAFTTIIIPAIIVGGIATGAVGTVSG 60  
QY 61 LSWGLKQAEAEANKTPDKPKVRIQAGKGFNEFPNKEYDLYRSLSSKIDGGWGNAA 120  
Db 61 LSWGLKQAEAEANKTPDKPKVRIQAGKGFNEFPNKEYDLYRSLSSKIDGGWGNAA 120  
QY 121 HYWYKGGQONKLEVDMDKDAVGT 143  
Db 121 HYWYKGGQONKLEVDMDKDAVGT 143

RESULT 3  
ID O34110 PRELIMINARY; PRT; 1295 AA.  
AC O34110;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F37;  
RX MEDLINE=97339580; PubMed=9196179;

RA Ito Y., Azuma T., Ito S., Miyaji H., Hirai M., Yamazaki Y., Sato F.,  
RA Kato T., Kohli Y., Kuriyama M.;  
RT "Analysis and typing of the vacA gene from caga-positive strains of  
RT Helicobacter pylori isolated in Japan.";  
RL J. Clin. Microbiol. 35:1710-1714(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F37;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT noncytotoxic Helicobacter pylori.";  
RL J. Infect. Dis. 178:1391-1398(1998).  
RN [3]  
RP SEQUENCE OF 638-752 FROM N.A.  
RC STRAIN=JK2-55;  
RA Yamazaki Y., Kodama T., Graham D.Y., Kita M., Imanishi J., Kashima K.;  
RT "vacA genotypes of Helicobacter pylori in relation to caga status,  
RT cytotoxin production, or clinical outcome.";  
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF071095; AAC77450.1;  
DR EMBL; AF042736; AAB99828.1;  
SQ SEQUENCE 1295 AA; 139808 MW; B4B7F7AEF7901CB8 CRC64;

Query Match 11.0%; Score 143; DB 2; Length 1295;  
Best Local Similarity 100.0%; Pred. No. 3.4e-139;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEIQOTHRKNRPLVSLALVGVLSITPQOASHAFTTIIIPAIIVGGIATGAVGTVSG 60  
Db 1 MEIQOTHRKNRPLVSLALVGVLSITPQOASHAFTTIIIPAIIVGGIATGAVGTVSG 60  
QY 61 LSWGLKQAEAEANKTPDKPKVRIQAGKGFNEFPNKEYDLYRSLSSKIDGGWGNAA 120  
Db 61 LSWGLKQAEAEANKTPDKPKVRIQAGKGFNEFPNKEYDLYRSLSSKIDGGWGNAA 120  
QY 121 HYWYKGGQONKLEVDMDKDAVGT 143  
Db 121 HYWYKGGQONKLEVDMDKDAVGT 143

RESULT 4  
ID O92HT1 PRELIMINARY; PRT; 1290 AA.  
AC O92HT1;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCCTC 11637;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT noncytotoxic Helicobacter pylori.";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049653; AAD04290.1;  
SQ SEQUENCE 1290 AA; 139551 MW; 3DCE42C519352541 CRC64;

Query Match 7.6%; Score 99; DB 2; Length 1290;  
Best Local Similarity 100.0%; Pred. No. 2.1e-93;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 VGNAAAMFNDIDSATGYKPLIKINSADLIKTEHVLLKAKIIGYVSTGTNGISN 756  
|||||  
Db 692 VGNAAAMFNDIDSATGYKPLIKINSADLIKTEHVLLKAKIIGYVSTGTNGISN 751  
|||||  
QY 757 VNLEEOFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 795  
|||||  
Db 752 VNLEEOFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 790  
|||||

RESULT 5  
O32663 PRELIMINARY; PRT; 160 AA.  
AC O32663;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP022;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63262; AAB61871.1; -  
FT NON\_TER 1 160  
FT NON\_TER 1 160  
SQ SEQUENCE 160 AA; 17717 MW; 4CD55D434235477B CRC64;

Query Match 7.4%; Score 96; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 4.3e-91;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 121 HYWVGQGNKLEVDKMDAVGTYTSLGRNFTGGDLVNMOKATRLRGFGNSFTSYKD 180  
|||||  
Db 65 HYWVGQGNKLEVDKMDAVGTYTSLGRNFTGGDLVNMOKATRLRGFGNSFTSYKD 124  
|||||  
QY 181 SADRTTRVDFNAKNSIDNFVEINRVGSGAGRKAS 216  
|||||  
Db 125 SADRTTRVDFNAKNSIDNFVEINRVGSGAGRKAS 160  
|||||

RESULT 6  
O32660 PRELIMINARY; PRT; 160 AA.  
AC O32660;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP013;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63259; AAB61868.1; -  
FT NON\_TER 1 160  
FT NON\_TER 1 160  
SQ SEQUENCE 160 AA; 17855 MW; 1C39344B40108A00 CRC64;

Query Match 7.3%; Score 94; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 5.2e-89;  
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 GFNEFPNKEYDLYRSLSSKIDGWDGNAARHYWVGQGNKLEVDKMDAVGTYTSLGL 148  
|||||  
Db 33 GFNEFPNKEYDLYRSLSSKIDGWDGNAARHYWVGQGNKLEVDKMDAVGTYTSLGL 92  
|||||  
QY 149 RNFTGGDLVNMOKATRLRGFGNSFTSYKDSA 182  
|||||  
Db 93 RNFTGGDLVNMOKATRLRGFGNSFTSYKDSA 126  
|||||

RESULT 7  
Q9R844 PRELIMINARY; PRT; 148 AA.  
AC Q9R844;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC6;  
RX MEDLINE=98445420; PubMed=9770535;  
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,  
Kunstmann E., Dyrek I., Achtman M.;  
RT "Free recombination within Helicobacter pylori."  
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).  
DR EMBL; AJ009437; CAA08712.1; -  
DR InterPro: IPR000285;  
DR ProDom: PD002768; -; 1.  
FT NON\_TER 1 148  
FT NON\_TER 1 148  
SQ SEQUENCE 148 AA; 16572 MW; 434CF44C0EEAC29B CRC64;

Query Match 7.1%; Score 92; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 5.8e-87;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 89 GFNEFPNKEYDLYRSLSSKIDGWDGNAARHYWVGQGNKLEVDKMDAVGTYTSLGL 148  
|||||  
Db 21 GFNEFPNKEYDLYRSLSSKIDGWDGNAARHYWVGQGNKLEVDKMDAVGTYTSLGL 80  
|||||  
QY 149 RNFTGGDLVNMOKATRLRGFGNSFTSYKD 180  
|||||  
Db 81 RNFTGGDLVNMOKATRLRGFGNSFTSYKD 112  
|||||

RESULT 8  
Q9LB57 PRELIMINARY; PRT; 143 AA.  
AC Q9LB57;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPN326;  
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
Rayat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori."  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050395; AAF26577.1; -

```

FT  NON_TER      1      1
FT  NON_TER     143     143
SQ  SEQUENCE    143 AA; 14814 MW;  AFEL3BE186B75884  CRC64;

Query Match
Best Local Similarity 100.0%; DB 2; Length 143;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  447 VENLTGNIIVDGPLRVNNOVGVALAGSANSFEKAGTDTKNGTATFNNDISLGRFVNLK 506
    |||||||
Db  54 VENLTGNIIVDGPLRVNNOVGVALAGSANSFEKAGTDTKNGTATFNNDISLGRFVNLK 113
    |||||||

QY  507 VDAHTANFKGIDTNGGFFNTLDFSGVTDK 535
Db  114 VDAHTANFKGIDTNGGFFNTLDFSGVTDK 142

RESULT 9
Q9LB53 PRELIMINARY; PRT; 120 AA.
AC  Q9LB53
DT  01-OCT-2000 (TremBLrel. 15, Created)
DT  01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT  01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE  VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN  VACA.
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC  Helicobacter.
OX  NCBI_TaxID=210;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=87A300;
RA  Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA  Raviat J.M., Xu G., Rappuoli R.;
RT  "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL  Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF050399; AAF6581.1; -.
FT  NON_TER      1      1
FT  NON_TER     120     120
SQ  SEQUENCE    120 AA; 13206 MW;  C79E258AAFFB70  CRC64;

Query Match
Best Local Similarity 100.0%; DB 2; Length 120;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  507 VDAHTANFKGIDTNGGFFNTLDFSGVTDKVNINKLITASTNVAVKFNINELIVKTNGIS 566
    |||||||
Db  1 VDAHTANFKGIDTNGGFFNTLDFSGVTDKVNINKLITASTNVAVKFNINELIVKTNGIS 60
    |||||||

QY  567 VGEYTHFSEDIGSQSRINVRLETGTRS 594
    |||||||
Db  61 VGEYTHFSEDIGSQSRINVRLETGTRS 88
    |||||||

RESULT 10
Q32671 PRELIMINARY; PRT; 160 AA.
AC  Q32671
DT  01-JAN-1998 (TremBLrel. 05, Created)
DT  01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT  01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE  VACUOLATING CYTOTOXIN (FRAGMENT).
GN  VACA.
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC  Helicobacter.
OX  NCBI_TaxID=210;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=HP053;

```

```

RA  Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL  Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U63270; AAB61879.1; -.
FT  NON_TER      1      1
FT  NON_TER     160     160
SQ  SEQUENCE    160 AA; 17832 MW;  FB3C47739B77FBF  CRC64;

Query Match
Best Local Similarity 100.0%; DB 2; Length 160;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  130 NKLEVDKMDAVGTYTILSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADTRTVD 189
    |||||||
Db  74 NKLEVDKMDAVGTYTILSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADTRTVD 133
    |||||||

QY  190 FNAKNISIDNFVEINNRVSGAGRKAS 216
    |||||||
Db  134 FNAKNISIDNFVEINNRVSGAGRKAS 160
    |||||||

RESULT 11
Q32684 PRELIMINARY; PRT; 160 AA.
AC  Q32684
DT  01-JAN-1998 (TremBLrel. 05, Created)
DT  01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT  01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE  VACUOLATING CYTOTOXIN (FRAGMENT).
GN  VACA.
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC  Helicobacter.
OX  NCBI_TaxID=210;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=HP158;
RA  Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL  Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U63284; AAB61893.1; -.
FT  NON_TER      1      1
FT  NON_TER     160     160
SQ  SEQUENCE    160 AA; 17830 MW;  AE218561309B1934  CRC64;

Query Match
Best Local Similarity 100.0%; DB 2; Length 160;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  130 NKLEVDKMDAVGTYTILSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADTRTVD 189
    |||||||
Db  74 NKLEVDKMDAVGTYTILSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADTRTVD 133
    |||||||

QY  190 FNAKNISIDNFVEINNRVSGAGRKAS 216
    |||||||
Db  134 FNAKNISIDNFVEINNRVSGAGRKAS 160
    |||||||

RESULT 12
Q86250 PRELIMINARY; PRT; 371 AA.
AC  Q86250
DT  01-NOV-1998 (TremBLrel. 08, Created)
DT  01-NOV-1998 (TremBLrel. 08, Last sequence update)
DT  01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE  VACA (FRAGMENT).
GN  VACA.
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC  Helicobacter.
OX  NCBI_TaxID=210;
RN  [1]
RP  SEQUENCE FROM N.A.

```



```
RC STRAIN=MZ19A;
RX MEDLINE=98180430; PubMed=9521135;
RA Han S.R., Schreiber H.J., Bhakdi S., Loos M., Maeurer M.J.;
RT "vaca" genotypes and genetic diversity in clinical isolates of
RT Helicobacter pylori.";
RL Clin. Diagn. Lab. Immunol. 5:139-145(1998);
DR EMBL; AJ006967; CAA07356.1; -;
FT NON_TER 1 371
FT NON_TER 371 371
SQ SEQUENCE 371 AA; 40021 MW; 38E7043AE7AC4564 CRC64;

Query Match
Best Local Similarity 100.0%; Score 87; DB 2; Length 371;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTGNTVVDGFLRVNNOVGGYALAGSSANFEFRAGTDTKNGTATFNNDISLGRFVNK 506
Db 23 VENLTGNTVVDGFLRVNNOVGGYALAGSSANFEFRAGTDTKNGTATFNNDISLGRFVNK 82

QY 507 VDAHTANFKGIDTGNNGGFNTLDFSGVT 533
Db 83 VDAHTANFKGIDTGNNGGFNTLDFSGVT 109

RESULT 13
Q9LB51
ID Q9LB51 PRELIMINARY; PRT; 120 AA.
AC Q9LB51;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE VACUOLATING CYTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=87A300;
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,
RA Rayat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050401; AAF26583.1; -;
FT NON_TER 1 120
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13408 MW; F15E1EE7AD115527 CRC64;

Query Match
Best Local Similarity 100.0%; Score 86; DB 2; Length 120;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 643 PWGTSKLMFNNTLQGNVMDYSQFSLTIQGFNNQGTINYLVRGKVFATLSVGNAAA 702
Db 35 PWGTSKLMFNNTLQGNVMDYSQFSLTIQGFNNQGTINYLVRGKVFATLSVGNAAA 94

QY 703 MFMNNDIDSATGFYKPLINKSAQDL 728
Db 95 MFMNNDIDSATGFYKPLINKSAQDL 120

RESULT 14
Q32669
ID Q32669 PRELIMINARY; PRT; 160 AA.
AC Q32669;
DT 01-JAN-1998 (TremBLrel. 05, Created)
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE VACUOLATING CYTOXIN (FRAGMENT).
GN VACA.
```

```
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP048;
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63268; AAB61877.1; -;
FT NON_TER 1 160
FT NON_TER 160 160
SQ SEQUENCE 160 AA; 17889 MW; B33B54874227C05A CRC64;

Query Match
Best Local Similarity 100.0%; Score 78; DB 2; Length 160;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 AVGTTLTSLGRNFTGGDLVDVNMOKATLRLGQFNGNSFTSYKDSADRTTRVDFNAKNISID 198
Db 83 AVGTTLTSLGRNFTGGDLVDVNMOKATLRLGQFNGNSFTSYKDSADRTTRVDFNAKNISID 142

QY 199 NFVEINNRVSGAGRKAS 216
Db 143 NFVEINNRVSGAGRKAS 160

RESULT 15
Q9R848
ID Q9R848 PRELIMINARY; PRT; 148 AA.
AC Q9R848;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE VACUOLATING CYTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC42;
RA MEDLINE=98445420; PubMed=9770535;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009430; CAA08705.1; -;
FT NON_TER 1 148
FT NON_TER 148 148
SQ SEQUENCE 148 AA; 16632 MW; F1A77FFFE1E0B5AF CRC64;

Query Match
Best Local Similarity 100.0%; Score 72; DB 2; Length 148;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSLGRNFTGGDLVDVNMOKATLRLGQFNGNSFTSYKDSADRTTRVDFNAKNISIDNFVEIN 204
Db 77 LSLGRNFTGGDLVDVNMOKATLRLGQFNGNSFTSYKDSADRTTRVDFNAKNISIDNFVEIN 136

QY 205 NRVGSGAGRKAS 216
Db 137 NRVGSGAGRKAS 148

RESULT 16
Q32661
ID Q32661 PRELIMINARY; PRT; 160 AA.
AC Q32661;
DT 01-JAN-1998 (TremBLrel. 05, Created)
```

DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)  
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP019;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63260; AAB61869.1; -  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17914 MW; COD3330C2357456B CRC64;

Query Match 5.6%; Score 72; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 4e-66;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 LSGLRNFTGGDLVNMOKATRLRGQFNCSFTSYKDSADTRTRVDFNAKNISIDNFVEIN 204  
|||||  
Db 89 LSGLRNFTGGDLVNMOKATRLRGQFNCSFTSYKDSADTRTRVDFNAKNISIDNFVEIN 148  
|||||  
Qy 205 NRVGSGAGRKAS 216  
|||||  
Db 149 NRVGSGAGRKAS 160

## RESULT 17

O32674 PRELIMINARY; PRT; 160 AA.  
ID O32674;  
AC O32674;  
DT 01-JAN-1998 (TremBLrel. 05, Created)  
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)  
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP060;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63273; AAB61882.1; -  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17857 MW; F7973BC7A055B8DF CRC64;

Query Match 5.6%; Score 72; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 4e-66;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 LSGLRNFTGGDLVNMOKATRLRGQFNCSFTSYKDSADTRTRVDFNAKNISIDNFVEIN 204  
|||||  
Db 89 LSGLRNFTGGDLVNMOKATRLRGQFNCSFTSYKDSADTRTRVDFNAKNISIDNFVEIN 148  
|||||  
Qy 205 NRVGSGAGRKAS 216  
|||||  
Db 149 NRVGSGAGRKAS 160

## RESULT 18

O9LB50 PRELIMINARY; PRT; 109 AA.  
ID O9LB50  
AC O9LB50;

DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=G104;  
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
RAYAT J.M., Xu G., Rappuoli R.;  
RL "Allelic variation of vacA gene in the Chinese Helicobacter pylori."  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050402; AAF26584.1; -  
FT NON\_TER 109  
FT NON\_TER 109  
SQ SEQUENCE 109 AA; 11808 MW; 9C4ABCACB1AB3AA1 CRC64;

Query Match 5.5%; Score 71; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 3.2e-65;  
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 SHAAFTTIIIPAIIVGGIATGATGTVSGLLSWGLKQAEANKTPDKPKVWRIQAGKGF 90  
|||||  
Db 31 SHAAFTTIIIPAIIVGGIATGATGTVSGLLSWGLKQAEANKTPDKPKVWRIQAGKGF 90  
|||||  
Qy 91 NEFPNKEYDLY 101  
|||||  
Db 91 NEFPNKEYDLY 101

## RESULT 19

Q9R847 PRELIMINARY; PRT; 148 AA.  
ID Q9R847  
AC Q9R847;  
DT 01-MAY-2000 (TremBLrel. 13, Created)  
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC47;  
RA MEDLINE=98445420; PubMed=9770535;  
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,  
Kunstmann E., Dyrek I., Achtman M.;  
RL "Free recombination within Helicobacter pylori."  
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).  
DR EMBL; AJ009432; CAA08707.1; -  
FT NON\_TER 1  
FT NON\_TER 148  
FT NON\_TER 148  
SQ SEQUENCE 148 AA; 16590 MW; 1AE8341399BA0370 CRC64;

Query Match 5.4%; Score 70; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 4.6e-64;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 LSGLRNFTGGDLVNMOKATRLRGQFNCSFTSYKDSADTRTRVDFNAKNISIDNFVEIN 204  
|||||  
Db 77 LSGLRNFTGGDLVNMOKATRLRGQFNCSFTSYKDSADTRTRVDFNAKNISIDNFVEIN 136  
|||||  
Qy 205 NRVGSGAGRK 214  
|||||  
Db 137 NRVGSGAGRK 146

```
RESULT 20
006021
ID 006021 PRELIMINARY: PRT: 1290 AA.
AC 006021;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43526;
RA Ogura K., Maeda S., Kanai F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001358; AAB53868.1;
SQ SEQUENCE 1290 AA; 139181 MW; 6254062581580578 CRC64;

Query Match
Best Local Similarity 5.2%; Score 68; DB 2; Length 1290;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 VGNAAAMFNNDIDTSATGYKPLIKNSAQDLIKNTHEVLLKAKIIGYGVNSTGTNGISN 756
DQ 692 VGNAAAMFNNDIDTSATGYKPLIKNSAQDLIKNTHEVLLKAKIIGYGVNSTGTNGISN 751
QY 757 VNLEQFK 764
DQ 752 VNLEQFK 759

RESULT 21
087018 PRELIMINARY: PRT: 1323 AA.
AC 087018;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE INACTIVE CYTOTOXIN.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95-54 (J128);
RA Pagliaccia C., Cover T., Rappuoli R., Telford J.L., Reytrat J.M.;
RT "The mform of the Helicobacter pylori cytotoxin has cell type
RT specific activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:0-0(1998).
DR EMBL; U95971; AAC25911.1;
SQ SEQUENCE 1323 AA; 142972 MW; E3A8B724C02FA2E3 CRC64;

Query Match
Best Local Similarity 5.1%; Score 66; DB 2; Length 1323;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 NKLEYDMKDAVGTYYLSGLRFTGGDLVNMOKATLRLGQNGNSFTSYKDSADRTTRVD 189
DQ 130 NKLEYDMKDAVGTYYLSGLRFTGGDLVNMOKATLRLGQNGNSFTSYKDSADRTTRVD 189
QY 190 FNAKNI 195
DQ 190 FNAKNI 195
```

```
RESULT 22
032672 PRELIMINARY: PRT: 160 AA.
AC 032672;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP054;
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63271; AAB61880.1;
FT NON_TER 1
FT NON_TER 160
SQ SEQUENCE 160 AA; 17810 MW; A39C099AD91D39C0 CRC64;

Query Match
Best Local Similarity 5.0%; Score 65; DB 2; Length 160;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 TGGDLVNMOKATLRLGQNGNSFTSYKDSADRTTRVDNFNAKNISIDNFVEINNRYGSGA 211
DQ 96 TGGDLVNMOKATLRLGQNGNSFTSYKDSADRTTRVDNFNAKNISIDNFVEINNRYGSGA 155
QY 212 GRKAS 216
DQ 156 GRKAS 160

RESULT 23
032681 PRELIMINARY: PRT: 160 AA.
AC 032681;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP112;
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63280; AAB61889.1;
FT NON_TER 1
FT NON_TER 160
SQ SEQUENCE 160 AA; 17788 MW; 0A535C6AE2447690 CRC64;

Query Match
Best Local Similarity 5.0%; Score 65; DB 2; Length 160;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 TGGDLVNMOKATLRLGQNGNSFTSYKDSADRTTRVDNFNAKNISIDNFVEINNRYGSGA 211
DQ 96 TGGDLVNMOKATLRLGQNGNSFTSYKDSADRTTRVDNFNAKNISIDNFVEINNRYGSGA 155
QY 212 GRKAS 216
DQ 156 GRKAS 160
```

```

RESULT 24
O32687 ID O32687 PRELIMINARY; PRT; 160 AA.
AC O32687;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP197;
RA Beech R.N., Gotte M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63287; AAB61896.1; -.
FT NON_TER 1
FT NON_TER 160
SQ SEQUENCE 160 AA; 17817 MW; 006E2E781C2E10EF CRC64;

Query Match 5.0%; Score 65; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 7.8e-59;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 152 TGGDLVNMQRKATLRGFGNGSFTSYKDSADRTVRVDFNARNISIDNVEINNRVGS 211
Db 96 TGGDLVNMQRKATLRGFGNGSFTSYKDSADRTVRVDFNARNISIDNVEINNRVGS 155

Oy 212 GRKAS 216
Db 156 GRKAS 160

RESULT 25
O9RE83 ID O9RE83 PRELIMINARY; PRT; 96 AA.
AC O9RE83;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NL4601;
RA van Doorn L.J.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NL4601;
RX MEDLINE=98371099; PubMed=9705399;
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,
RA Atherton J.C., Blaser M.J., Quint W.G.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390596; CAB64366.1; -.
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10224 MW; 2F19D55DECB09BB9 CRC64;

Query Match 4.6%; Score 60; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 7.9e-54;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 535 KVNINKLITASTNAVKNFNINELIVKTNIGISVGEYTHFSEDIGSQSRINTVRLTGTRS 594
Db 29 KVNINKLITASTNAVKNFNINELIVKTNIGISVGEYTHFSEDIGSQSRINTVRLTGTRS 88

RESULT 26
O9R842 ID O9R842 PRELIMINARY; PRT; 148 AA.
AC O9R842;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FA3;
RX MEDLINE=98445420; PubMed=9770535;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009439; CAA08714.1; -.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16629 MW; 1CC7FEBB3DC71F7 CRC64;

Query Match 4.6%; Score 60; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.2e-53;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 69 EEANKTPDKPDKVRIQAGKGFNEFPNKEYDLYRSLSSKIDGWDGNAARHYVWKGQ 128
Db 1 EEANKTPDKPDKVRIQAGKGFNEFPNKEYDLYRSLSSKIDGWDGNAARHYVWKGQ 60

RESULT 27
O9KX03 ID O9KX03 PRELIMINARY; PRT; 207 AA.
AC O9KX03;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=003;
RA Bereswill S.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=003;
RX MEDLINE=98233879; PubMed=9574692;
RA Strobel S., Bereswill S., Allgauer P., Ballig P., Sonntag H.G.,
RA Kist M.;
RT "Identification and analysis of a new vacA genotype variant of
RT Helicobacter pylori in different patient groups in Germany.";
RL J. Clin. Microbiol. 36:1285-1289(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=003;
RA Strobel S.;

```

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y14740; CAB69819.1; -  
 FT NON\_TER 1  
 FT NON\_TER 207  
 SQ SEQUENCE 207 AA; 22923 MW; 04535F329C2E9EC7 CRC64;

Query Match 4.6%; Score 60; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-53;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNWYFDARNIKNVEITNKLAFGPGSPWGTSKLMFNLTGQNAVMDYSQFSNLTIOGDF 676  
 Db 51 PNWYFDARNIKNVEITNKLAFGPGSPWGTSKLMFNLTGQNAVMDYSQFSNLTIOGDF 110

RESULT 28

Q9KX02 ID Q9KX02 PRELIMINARY; PRT; 207 AA.  
 AC Q9KX02;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=013;  
 RA Strobel S.;  
 RA Bereswill S.;  
 RA Kist M.;  
 RT Identification and analysis of a new vacA genotype variant of  
 RT Helicobacter pylori in different patient groups in Germany.;  
 RL J. Clin. Microbiol. 36:1285-1289(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=013;  
 RX MEDLINE=98233879; PubMed=9574692;  
 RA Strobel S.; Bereswill S.; Allgaier P.; Balig P.; Sonntag H.G.;  
 RA Kist M.;  
 RT Identification and analysis of a new vacA genotype variant of  
 RT Helicobacter pylori in different patient groups in Germany.;  
 RL J. Clin. Microbiol. 36:1285-1289(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=013;  
 RA Strobel S.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y14741; CAB69817.1; -  
 FT NON\_TER 1  
 FT NON\_TER 207  
 SQ SEQUENCE 207 AA; 22962 MW; 5BA1260D96FE6218 CRC64;

Query Match 4.6%; Score 60; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-53;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNWYFDARNIKNVEITNKLAFGPGSPWGTSKLMFNLTGQNAVMDYSQFSNLTIOGDF 676  
 Db 51 PNWYFDARNIKNVEITNKLAFGPGSPWGTSKLMFNLTGQNAVMDYSQFSNLTIOGDF 110

RESULT 29

Q9KX01 ID Q9KX01 PRELIMINARY; PRT; 207 AA.  
 AC Q9KX01;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).

Query Match 4.6%; Score 60; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-53;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNWYFDARNIKNVEITNKLAFGPGSPWGTSKLMFNLTGQNAVMDYSQFSNLTIOGDF 676  
 Db 51 PNWYFDARNIKNVEITNKLAFGPGSPWGTSKLMFNLTGQNAVMDYSQFSNLTIOGDF 110

RESULT 30

Q9R305 ID Q9R305 PRELIMINARY; PRT; 96 AA.  
 AC Q9R305;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACA PROTEIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CR18;  
 RA van Doorn L.J.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CR18;  
 RX MEDLINE=98371099; PubMed=9705399;  
 RA van Doorn L.J.; Figueiredo C.; Sanna R.; Pena S.; Midolo P.; Ng E.K.;  
 RA Atherton J.C.; Blaser M.J.; Quint W.G.;  
 RT "Expanding allelic diversity of Helicobacter pylori vacA";  
 RL J. Clin. Microbiol. 36:2597-2603(1998).  
 DR EMBL; AJ390606; CAB64281.1; -  
 DR EMBL; AJ390605; CAB64280.1; -  
 FT NON\_TER 1  
 FT NON\_TER 96  
 SQ SEQUENCE 96 AA; 10268 MW; 40D6D55DEE71F380 CRC64;

Query Match 4.5%; Score 58; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-52;  
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=026;  
 RA Bereswill S.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=026;  
 RX MEDLINE=98233879; PubMed=9574692;  
 RA Strobel S.; Bereswill S.; Allgaier P.; Balig P.; Sonntag H.G.;  
 RA Kist M.;  
 RT Identification and analysis of a new vacA genotype variant of  
 RT Helicobacter pylori in different patient groups in Germany.;  
 RL J. Clin. Microbiol. 36:1285-1289(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=026;  
 RA Strobel S.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y14742; CAB69818.1; -  
 FT NON\_TER 1  
 FT NON\_TER 207  
 SQ SEQUENCE 207 AA; 23000 MW; D69B44B7D46C2AED CRC64;

Query Match 4.6%; Score 60; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-53;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNWYFDARNIKNVEITNKLAFGPGSPWGTSKLMFNLTGQNAVMDYSQFSNLTIOGDF 676  
 Db 51 PNWYFDARNIKNVEITNKLAFGPGSPWGTSKLMFNLTGQNAVMDYSQFSNLTIOGDF 110

RESULT 30

Q9R305 ID Q9R305 PRELIMINARY; PRT; 96 AA.  
 AC Q9R305;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACA PROTEIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CR18;  
 RA van Doorn L.J.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CR18;  
 RX MEDLINE=98371099; PubMed=9705399;  
 RA van Doorn L.J.; Figueiredo C.; Sanna R.; Pena S.; Midolo P.; Ng E.K.;  
 RA Atherton J.C.; Blaser M.J.; Quint W.G.;  
 RT "Expanding allelic diversity of Helicobacter pylori vacA";  
 RL J. Clin. Microbiol. 36:2597-2603(1998).  
 DR EMBL; AJ390606; CAB64281.1; -  
 DR EMBL; AJ390605; CAB64280.1; -  
 FT NON\_TER 1  
 FT NON\_TER 96  
 SQ SEQUENCE 96 AA; 10268 MW; 40D6D55DEE71F380 CRC64;

Query Match 4.5%; Score 58; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-52;  
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 507 VDAHTANFKGIDFGNGFNTLDFSGVTDKVNINKLITASTNNVAVKFNINELIVKTN 564
|||||
Db 1 VDAHTANFKGIDFGNGFNTLDFSGVTDKVNINKLITASTNNVAVKFNINELIVKTN 58
|||||

RESULT 31
Q9LBA7
ID Q9LBA7 PRELIMINARY; PRT; 91 AA.
AC Q9LBA7
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN3342A;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RA "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030343; AAF26525.1; -.
FT NON_TER 91
SQ SEQUENCE 91 AA; 9737 MW; 0D1A97E81C4D3321 CRC64;

Query Match 4.4%; Score 57; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 9.9e-51;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 SHAFFTTVIPAIVGGIATGATGVTGSLLSWGLKQAEANKTPDKPKVWRIQAG 87
|||||
Db 31 SHAFFTTVIPAIVGGIATGATGVTGSLLSWGLKQAEANKTPDKPKVWRIQAG 87
|||||

RESULT 32
Q32677
ID Q32677 PRELIMINARY; PRT; 160 AA.
AC Q32677;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP071;
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63276; AAB61885.1; -.
FT NON_TER 160
SQ SEQUENCE 160 AA; 17715 MW; 6A015EDA5338EE26 CRC64;

Query Match 4.4%; Score 57; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.6e-50;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 WGLKQAEANKTPDKPKVWRIQAGKGFNEFPNKEYDYLRSLSSKIDGGWDGNA 119
|||||
Db 7 WGLKQAEANKTPDKPKVWRIQAGKGFNEFPNKEYDYLRSLSSKIDGGWDGNA 63
|||||
```

```
RESULT 33
Q9R313
ID Q9R313 PRELIMINARY; PRT; 148 AA.
AC Q9R313;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC57, AND CC24;
RX MEDLINE=98445420; PubMed=9770535;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009435; CAA08710.1; -.
DR EMBL; AJ009418; CAA08693.1; -.
FT NON_TER 148
SQ SEQUENCE 148 AA; 16591 MW; 6E38DA05BE454A13 CRC64;

Query Match 4.3%; Score 56; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e-49;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFTGGDLVNMOKATRLRGQFNNGSFTSYKDSADRTTRVDFNAKNISIDNF 200
|||||
Db 77 LSGLRNFTGGDLVNMOKATRLRGQFNNGSFTSYKDSADRTTRVDFNAKNISIDNF 132
|||||

RESULT 34
Q32683
ID Q32683 PRELIMINARY; PRT; 160 AA.
AC Q32683;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP151;
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63282; AAB61891.1; -.
FT NON_TER 160
SQ SEQUENCE 160 AA; 17872 MW; 303AFCAAC055B5AE CRC64;

Query Match 4.3%; Score 56; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.8e-49;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFTGGDLVNMOKATRLRGQFNNGSFTSYKDSADRTTRVDFNAKNISIDNF 200
|||||
Db 89 LSGLRNFTGGDLVNMOKATRLRGQFNNGSFTSYKDSADRTTRVDFNAKNISIDNF 144
|||||

RESULT 35
Q9F9G3
ID Q9F9G3 PRELIMINARY; PRT; 862 AA.
```

```

AC Q9F9G3;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DE 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=62;
RA Kuo C.H., Wang W.C.;
RT "Genetic and functional comparison of Helicobacter pylori vacuolating
RL toxin between sla/mit and sla/m2 isolates.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF195012; AG28440.1; -.
FT NON_TER 1
FT NON_TER 862
SQ SEQUENCE 862 AA; 92778 MW; 42D90931B97C09B3 CRC64;

Query Match 4.3%; Score 56; DB 2; Length 862;
Best Local Similarity 100.0%; Pred. No. 8.1e-49;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 761 EGFKERLALYNNRMDTCVVRNTDDIKACGMAIGDQSMVNNPNKYLYLKAWKN 816
DB 753 EGFKERLALYNNRMDTCVVRNTDDIKACGMAIGDQSMVNNPNKYLYLKAWKN 808

RESULT 36
Q9R851
ID Q9R851 PRELIMINARY; PRT; 148 AA.
AC Q9R851;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DE 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC2;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009423; CA008698.1; -.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16644 MW; 8868454CE555213E CRC64;

Query Match 4.2%; Score 55; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.8e-48;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 GFNEFPNKEYDLYRSLSSKIDGWDGNGNARHYWVGGQONKLEVDKDAVGTY 143
DB 21 GFNEFPNKEYDLYRSLSSKIDGWDGNGNARHYWVGGQONKLEVDKDAVGTY 75

RESULT 37
Q9L811
ID Q9L811 PRELIMINARY; PRT; 78 AA.
AC Q9L811;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INDIA10;
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,
RA Nair G.B., Berg D.E.;
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,
RL India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL; AF217729; AAF42840.1; -.
FT NON_TER 1
FT NON_TER 78
SQ SEQUENCE 78 AA; 8077 MW; B578E3141735602D CRC64;

Query Match 4.0%; Score 52; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.4e-45;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 INRPLVSLALVGLVSIPTQOSHAAFFTTVIIPAIVGGIATGTAAGTVSGLL 61
DB 2 INRPLVSLALVGLVSIPTQOSHAAFFTTVIIPAIVGGIATGTAAGTVSGLL 53

RESULT 38
Q9L807
ID Q9L807 PRELIMINARY; PRT; 78 AA.
AC Q9L807;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DE 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INDIA67;
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,
RA Nair G.B., Berg D.E.;
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,
RL India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL; AF217735; AAF42846.1; -.
FT NON_TER 1
FT NON_TER 78
SQ SEQUENCE 78 AA; 8104 MW; B3C8E5A41731357C CRC64;

Query Match 4.0%; Score 52; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.4e-45;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 INRPLVSLALVGLVSIPTQOSHAAFFTTVIIPAIVGGIATGTAAGTVSGLL 61
DB 2 INRPLVSLALVGLVSIPTQOSHAAFFTTVIIPAIVGGIATGTAAGTVSGLL 53

RESULT 39
Q9K2T9

```

ID Q9K2T9 PRELIMINARY; PRT; 78 AA.  
AC Q9K2T9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-INDIA18, INDIA9, AND INDIA17;  
RX MEDLINE=20270153; PubMed=10809703;  
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,  
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,  
RA Nair G.B., Berg D.E.;  
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,  
RT India";  
RL J. Bacteriol. 182:3219-3227(2000).  
DR EMBL; AF217731; AAF42842.1; -;  
DR EMBL; AF217728; AAF42839.1; -;  
DR EMBL; AF217730; AAF42841.1; -;  
DR InterPro: IPR000755; -;  
DR ProDom: PD013478; -; 1.  
FT NON\_TER 1 1  
FT NON\_TER 78 78  
SQ SEQUENCE 78 AA; 8089 MW; A0B70924172F5304 CRC64;

Query Match 4.0%; Score 52; DB 2; Length 78;  
Best Local Similarity 100.0%; Pred. No. 1.4e-45;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 INRPLVSLALVGVLSITPQSHAAFFTTIIPAVGGIATGAVTVSGLL 61  
|||||  
DB 2 INRPLVSLALVGVLSITPQSHAAFFTTIIPAVGGIATGAVTVSGLL 53

RESULT 40

OY Q9RE76 PRELIMINARY; PRT; 96 AA.  
AC Q9RE76;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE VACA PROTEIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AU10;  
RA van Doorn L.J.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AU10;  
RX MEDLINE=98371099; PubMed=9705399;  
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,  
RA Atherton J.C., Blaser M.J., Quint W.G.;  
RT "Expanding allelic diversity of Helicobacter pylori vacA";  
RL J. Clin. Microbiol. 36:2597-2603(1998).  
DR EMBL; AJ390610; CAB64285.1; -;  
FT NON\_TER 1 1  
FT NON\_TER 96 96  
SQ SEQUENCE 96 AA; 10235 MW; 40D6D544D6090080 CRC64;

Query Match 4.0%; Score 52; DB 2; Length 96;

Best Local Similarity 100.0%; Pred. No. 1.7e-45;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 507 VDAHTANFKGIDTGGGNTLDFSGVTDKVNINKLITASTNVAVKNFINEL 558  
|||||  
DB 1 VDAHTANFKGIDTGGGNTLDFSGVTDKVNINKLITASTNVAVKNFINEL 52

RESULT 41

OY Q9ZHU8 PRELIMINARY; PRT; 1291 AA.  
AC Q9ZHU8;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-F52;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT nontoxic Helicobacter pylori";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049631; AAD04269.1; -;  
SQ SEQUENCE 1291 AA; 139473 MW; 398037E9EF290254 CRC64;

Query Match 4.0%; Score 52; DB 2; Length 1291;  
Best Local Similarity 100.0%; Pred. No. 1.7e-44;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1193 SSOHLFNASANVEARYYYGDTSYFYMNAGVLQEFARHVGNSNAASLNTFKVNA 1244  
|||||  
DB 1188 SSOHLFNASANVEARYYYGDTSYFYMNAGVLQEFARHVGNSNAASLNTFKVNA 1239

RESULT 42

OY Q9ZHU7 PRELIMINARY; PRT; 1291 AA.  
AC Q9ZHU7;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-F55;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT nontoxic Helicobacter pylori";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049632; AAD04270.1; -;  
SQ SEQUENCE 1291 AA; 139347 MW; 0452A0F725A18672 CRC64;

Query Match 4.0%; Score 52; DB 2; Length 1291;  
Best Local Similarity 100.0%; Pred. No. 1.7e-44;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1193 SSQHLFNASANVEARYYYGDTSYFYMNAGVLOEFAHVGSGNNAASLNTFKVNA 1244  
 |||||  
 Db 1188 SSQHLFNASANVEARYYYGDTSYFYMNAGVLOEFAHVGSGNNAASLNTFKVNA 1239

## RESULT 43

Q9LB54  
 ID Q9LB54 PRELIMINARY; PRT; 114 AA.  
 AC Q9LB54;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=87A300;  
 RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,  
 RA Rayrat J.M., Xu G., Rappuoli R.;  
 RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF050398; AAF26580.1; -  
 FT NON\_TER 1  
 FT NON\_TER 114  
 SQ SEQUENCE 114 AA; 12356 MW; C5589927F8B69CFF CRC64;

Query Match 3.9%; Score 51; DB 2; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-44;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 686 LVGRGKVPATLSVGNAAAMFNNDISATGFKYPLKINSADLIKNTEHVL 736  
 |||||  
 Db 1 LVGRGKVPATLSVGNAAAMFNNDISATGFKYPLKINSADLIKNTEHVL 51

## RESULT 44

Q9LB56  
 ID Q9LB56 PRELIMINARY; PRT; 143 AA.  
 AC Q9LB56;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPM326;  
 RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,  
 RA Rayrat J.M., Xu G., Rappuoli R.;  
 RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF050396; AAF26578.1; -  
 FT NON\_TER 1  
 FT NON\_TER 143  
 SQ SEQUENCE 143 AA; 15679 MW; F490221919C1A7F4 CRC64;

Query Match 3.9%; Score 51; DB 2; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-44;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 VGNAAAMFNNDISATGFKYPLKINSADLIKNTEHVLKAKIYGYNV 747  
 |||||  
 Db 41 VGNAAAMFNNDISATGFKYPLKINSADLIKNTEHVLKAKIYGYNV 91

## RESULT 45

Q9LB48  
 ID Q9LB48 PRELIMINARY; PRT; 145 AA.  
 AC Q9LB48;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=G104;  
 RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,  
 RA Rayrat J.M., Xu G., Rappuoli R.;  
 RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF050404; AAF26586.1; -  
 FT NON\_TER 1  
 FT NON\_TER 145  
 SQ SEQUENCE 145 AA; 15912 MW; 66FF5FE7F2CF61F1 CRC64;

Query Match 3.9%; Score 51; DB 2; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-44;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 749 TGTGINSVNLEEOFKERLALYNNNNRMDTCVVRNTDDIKACGMAIGDSM 799  
 |||||  
 Db 95 TGTGINSVNLEEOFKERLALYNNNNRMDTCVVRNTDDIKACGMAIGDSM 145

## RESULT 46

Q9S316  
 ID Q9S316 PRELIMINARY; PRT; 148 AA.  
 AC Q9S316;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=#12;  
 RA MEDLINE=99255683; PubMed=10320570;  
 RA Achtmann M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,  
 RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;  
 RT "Recombination and clonal groupings within Helicobacter pylori from  
 different geographic regions.";  
 RL Mol. Microbiol. 32:459-470(1999).  
 DR EMBL; AJ239550; CAB37673.1; -  
 FT NON\_TER 1  
 FT NON\_TER 148  
 SQ SEQUENCE 148 AA; 16516 MW; 7498EB67E37CB953 CRC64;

Query Match 3.9%; Score 51; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-44;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EANKTPDKPKVWRIQAGKGFNEFPNKEYDLYRSLSSKIDGGWGNAA 119  
 |||||  
 Db 1 EANKTPDKPKVWRIQAGKGFNEFPNKEYDLYRSLSSKIDGGWGNAA 51

```

RESULT 47
Q9S3I3
ID Q9S3I3 PRELIMINARY; PRT; 148 AA.
AC Q9S3I3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96-228;
RX MEDLINE=99255683; PubMed=10320570;
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239558; CAB37681.1; -.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16672 MW; 2ACA6241DCE1C458 CRC64;

```

```

Query Match 3.9%; Score 51; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.7e-44;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 LSGLNFTGGDLVNMOKATLRLGQFNNGNSFTSYKDSADTRTRVDFNAKNI 195
|||||
Db 77 LSGLNFTGGDLVNMOKATLRLGQFNNGNSFTSYKDSADTRTRVDFNAKNI 127
|||||

```

```

RESULT 48
Q9S3I1
ID Q9S3I1 PRELIMINARY; PRT; 148 AA.
AC Q9S3I1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N6;
RX MEDLINE=99255683; PubMed=10320570;
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239561; CAB37373.1; -.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16603 MW; E45B981AE8CC04F CRC64;

```

```

Query Match 3.9%; Score 51; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.7e-44;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 NKLEVDKMDAVGTTLSGLRNFTHGDLVNMOKATLRLGQFNNGNSFTSYKD 180
|||||
Db 62 NKLEVDKMDAVGTTLSGLRNFTHGDLVNMOKATLRLGQFNNGNSFTSYKD 112
|||||

```

```

RESULT 49
Q9S3H8
ID Q9S3H8 PRELIMINARY; PRT; 148 AA.
AC Q9S3H8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96-232;
RX MEDLINE=99255683; PubMed=10320570;
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239567; CAB37379.1; -.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16617 MW; 6E38DA05A8299A13 CRC64;

```

```

Query Match 3.9%; Score 51; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.7e-44;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 EEANKTPDKPWRIQAGKGFNEFPNKEYDLYRLSSKIDGGDWGNAA 119
|||||
Db 1 EEANKTPDKPWRIQAGKGFNEFPNKEYDLYRLSSKIDGGDWGNAA 51
|||||

```

```

RESULT 50
Q32664
ID Q32664 PRELIMINARY; PRT; 160 AA.
AC Q32664;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP024;
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63263; AAB61672.1; -.
FT NON_TER 1
FT NON_TER 160
SQ SEQUENCE 160 AA; 17872 MW; EE75577A191DC37F CRC64;

```

```

Query Match 3.9%; Score 51; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 2.9e-44;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 LSGLNFTGGDLVNMOKATLRLGQFNNGNSFTSYKDSADTRTRVDFNAKNI 195
|||||
Db 89 LSGLNFTGGDLVNMOKATLRLGQFNNGNSFTSYKDSADTRTRVDFNAKNI 139
|||||

```

```

RESULT 51

```

Q9XX00  
ID Q9XX00 PRELIMINARY; PRT; 207 AA.  
AC Q9XX00;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=031;  
RA Bereswill S.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=031;  
RA MEDLINE=98233879; PubMed=9574692;  
RA Strobel S.;  
RA Kist M.;  
RT Identification and analysis of a new vacA genotype variant of  
RT Helicobacter pylori in different patient groups in Germany.;  
RL J. Clin. Microbiol. 36:1285-1289(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=031;  
RA Strobel S.;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y14743; CAB69820.1;  
FT NON\_TER 1  
FT NON\_TER 207  
SQ SEQUENCE 207 AA; 22903 MW; 0FCCC71E0DD081C1 CRC64;

Query Match 3.9%; Score 51; DB 2; Length 207;  
Best Local Similarity 100.0%; Pred. No. 3.6e-44;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 640 QGSPWGTSKLMFNNTLQGNVMDYSQFSNLTQGFNNQGTINLYVRGG 690  
Db 74 QGSPWGTSKLMFNNTLQGNVMDYSQFSNLTQGFNNQGTINLYVRGG 124

RESULT 52  
O32670  
ID O32670 PRELIMINARY; PRT; 160 AA.  
AC O32670;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP051;  
RA Beech R.N.; Gottke M.U.; Fallone C.A.; Loo V.; Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63269; AAB61878.1;  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17830 MW; 9CEB57BEE1BDF359 CRC64;

Query Match 3.9%; Score 50; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 3.2e-43;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 AVGYTTLGSLRNFTGDLVDVNMOKATRLRGQFNNGNSFTSYKDSADRTTRV 188  
Db 83 AVGYTTLGSLRNFTGDLVDVNMOKATRLRGQFNNGNSFTSYKDSADRTTRV 132

RESULT 53  
O9LB81  
ID O9LB81 PRELIMINARY; PRT; 85 AA.  
AC O9LB81;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CHN4611A;  
RA Ji X.; Telford J.L.; Burroni D.; Guidotti S.; Pagliaccia C.;  
RA Rayrat J.M.; Xu G.; Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori."  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050369; AAF26531.1;  
FT NON\_TER 85  
SQ SEQUENCE 85 AA; 8999 MW; 882562CAC29F4DC9 CRC64;

Query Match 3.8%; Score 49; DB 2; Length 85;  
Best Local Similarity 100.0%; Pred. No. 2e-42;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIQOQTHRKINRPLVSLVGVLSITPQOSHAAFTTIIIPAIYGGIA 49  
Db 1 MEIQOQTHRKINRPLVSLVGVLSITPQOSHAAFTTIIIPAIYGGIA 49

RESULT 54  
O9LB95  
ID O9LB95 PRELIMINARY; PRT; 98 AA.  
AC O9LB95;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CHN3948A;  
RA Ji X.; Telford J.L.; Burroni D.; Guidotti S.; Pagliaccia C.;  
RA Rayrat J.M.; Xu G.; Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori."  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050355; AAF26537.1;  
FT NON\_TER 1  
FT NON\_TER 98  
SQ SEQUENCE 98 AA; 10373 MW; 93C668F44CE1B74F CRC64;

Query Match 3.7%; Score 48; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.5e-41;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTDLWQSGAGLNTIAPPEGYKDKPNNTPSQSGAKNDKNSAKNDKQ 357  
Db 40 IGTDLWQSGAGLNTIAPPEGYKDKPNNTPSQSGAKNDKNSAKNDKQ 87

```

RESULT 55
Q9LBA6 PRELIMINARY; PRT; 130 AA.
ID Q9LBA6
AC Q9LBA6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN3342A;
RA Ji X., Telford J.L., Burrioni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050344; AAF26526.1; -
FT NON_TER 1
FT NON_TER 130
SQ SEQUENCE 130 AA; 13776 MW; EA35E869AD6EFB52 CRC64;

Query Match 3.7%; Score 48; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.2e-41;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 IGTLDLWQSLNIIAPPEGGYKDKPNTPSQGAKNDKNSAKNDKQ 357
Db 40 IGTLDLWQSLNIIAPPEGGYKDKPNTPSQGAKNDKNSAKNDKQ 87

RESULT 56
Q32676 PRELIMINARY; PRT; 160 AA.
ID Q32676
AC Q32676;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP067;
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63275; AAB61884.1; -
FT NON_TER 1
FT NON_TER 160
SQ SEQUENCE 160 AA; 17858 MW; F7793BC7A055B8D5 CRC64;

Query Match 3.7%; Score 48; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 3.9e-41;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 LSGLRNFTGGDLVNMOKATLRLQFNGNSFTSKYKDSADTRTRVDNA 192
Db 89 LSGLRNFTGGDLVNMOKATLRLQFNGNSFTSKYKDSADTRTRVDNA 136

RESULT 57
Q9R958 PRELIMINARY; PRT; 1296 AA.
ID Q9R958

```

```

Q9R958;
AC 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F73;
RX MEDLINE=98453456; PubMed=9780260;
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
RA Kuriyama M.;
RT "Full-length sequence analysis of the vacA gene from cytotoxic and
RT noncytotoxic Helicobacter pylori.";
RL J. Infect. Dis. 178:1391-1398(1998).
DR EMBL; AF049652; AAD04289.1; -
SQ SEQUENCE 1296 AA; 139862 MW; 72469BEE8000BFDA CRC64;

Query Match 3.7%; Score 48; DB 2; Length 1296;
Best Local Similarity 100.0%; Pred. No. 2.4e-40;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 QOTHRKNRPLVSLALVGALVSITPQSHAAFFTTVIPAIVGGIATG 51
Db 4 QOTHRKNRPLVSLALVGALVSITPQSHAAFFTTVIPAIVGGIATG 51

RESULT 58
Q9LBC7 PRELIMINARY; PRT; 1324 AA.
ID Q9LBC7
AC Q9LBC7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN5147C;
RA Ji X., Telford J.L., Burrioni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050320; AAF26502.1; -
SQ SEQUENCE 1324 AA; 143414 MW; 739AF6E0C05C1D2 CRC64;

Query Match 3.7%; Score 48; DB 2; Length 1324;
Best Local Similarity 100.0%; Pred. No. 2.5e-40;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 QOTHRKNRPLVSLALVGALVSITPQSHAAFFTTVIPAIVGGIATG 51
Db 4 QOTHRKNRPLVSLALVGALVSITPQSHAAFFTTVIPAIVGGIATG 51

RESULT 59
Q9LBC3 PRELIMINARY; PRT; 1324 AA.
ID Q9LBC3
AC Q9LBC3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).

```

```
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN1811A;
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050326; AAF26508.1; -
FT NON_TER. 1324 1324
SQ SEQUENCE 1324 AA; 143271 MW; 661AE59F64C368C CRC64;

Query Match 3.7%; Score 48; DB 2; Length 1324;
Best Local Similarity 100.0%; Pred. No. 2.5e-40;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 QOHRKINPLVSLALVGVLSITPQSHAAFTTIIIPAIYGGIATG 51
DB 4 QOHRKINPLVSLALVGVLSITPQSHAAFTTIIIPAIYGGIATG 51

RESULT 60
Q9LBC2
ID Q9LBC2 PRELIMINARY; PRT; 1324 AA.
AC Q9LBC2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN5114A;
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050327; AAF26509.1; -
SQ SEQUENCE 1324 AA; 143288 MW; E4E516A5330CAB14 CRC64;

Query Match 3.7%; Score 48; DB 2; Length 1324;
Best Local Similarity 100.0%; Pred. No. 2.5e-40;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 QOHRKINPLVSLALVGVLSITPQSHAAFTTIIIPAIYGGIATG 51
DB 4 QOHRKINPLVSLALVGVLSITPQSHAAFTTIIIPAIYGGIATG 51

RESULT 61
Q9LB41
ID Q9LB41 PRELIMINARY; PRT; 119 AA.
AC Q9LB41;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=G39;
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050411; AAF26593.1; -
FT NON_TER. 119 119
SQ SEQUENCE 119 AA; 13034 MW; C56A6E477BC34C83 CRC64;

Query Match 3.6%; Score 47; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 3.2e-40;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 749 TGTNGISNVNLEEQKFERLALYNNNNRMDTCVVRNTDDIKACGMAIG 795
DB 71 TGTNGISNVNLEEQKFERLALYNNNNRMDTCVVRNTDDIKACGMAIG 117

RESULT 62
Q9ZHT2
ID Q9ZHT2 PRELIMINARY; PRT; 1291 AA.
AC Q9ZHT2;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F72;
RX MEDLINE=98453456; PubMed=9780260;
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
RA Kuriyama M.;
RT "Full-length sequence analysis of the vacA gene from cytotoxic and
noncytotoxic Helicobacter pylori.";
RJ Infect. Dis. 178:1391-1398(1998).
DR EMBL; AF049651; AAD04288.1; -
SQ SEQUENCE 1291 AA; 139720 MW; D3FC79263B4B1069 CRC64;

Query Match 3.6%; Score 47; DB 2; Length 1291;
Best Local Similarity 100.0%; Pred. No. 2.7e-39;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 EINNRVSGAGRKASSTVLTQASEGITSKNAEISLYDGATNLAS 248
DB 202 EINNRVSGAGRKASSTVLTQASEGITSKNAEISLYDGATNLAS 248

RESULT 63
Q9ZHV2
ID Q9ZHV2 PRELIMINARY; PRT; 1296 AA.
AC Q9ZHV2;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F35;
RX MEDLINE=98453456; PubMed=9780260;
```

RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT nontoxic Helicobacter pylori.";  
J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049625; AAD04264.1;  
SQ SEQUENCE 1296 AA; 139900 MW; A5F86C6BF6C37C9C CRC64;

Query Match 3.6%; Score 47; DB 2; Length 1296;  
Best Local Similarity 100.0%; Pred. No. 2.7e-39;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 EINNVRGSGAGRKASSTVLTLOASEGITSKNAEISLYDGATLNLAS 248  
Db 202 EINNVRGSGAGRKASSTVLTLOASEGITSKNAEISLYDGATLNLAS 248

RESULT 64  
Q9ZHV1 PRELIMINARY; PRT; 1296 AA.  
ID Q9ZHV1;  
AC Q9ZHV1;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT nontoxic Helicobacter pylori.";  
J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049626; AAD04265.1;  
SQ SEQUENCE 1296 AA; 139991 MW; AB9E313B29817EFF CRC64;

Query Match 3.6%; Score 47; DB 2; Length 1296;  
Best Local Similarity 100.0%; Pred. No. 2.7e-39;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 EINNVRGSGAGRKASSTVLTLOASEGITSKNAEISLYDGATLNLAS 248  
Db 202 EINNVRGSGAGRKASSTVLTLOASEGITSKNAEISLYDGATLNLAS 248

RESULT 65  
Q9LBC8 PRELIMINARY; PRT; 1328 AA.  
ID Q9LBC8;  
AC Q9LBC8;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CHN3295B;  
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF050319; AAF26501.1;  
SQ SEQUENCE 1328 AA; 143701 MW; 78C2298479CDEF3A CRC64;

Query Match 3.6%; Score 47; DB 2; Length 1328;  
Best Local Similarity 100.0%; Pred. No. 2.7e-39;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 EINNVRGSGAGRKASSTVLTLOASEGITSKNAEISLYDGATLNLAS 248  
Db 202 EINNVRGSGAGRKASSTVLTLOASEGITSKNAEISLYDGATLNLAS 248

RESULT 66  
Q9S310 PRELIMINARY; PRT; 148 AA.  
ID Q9S310;  
AC Q9S310;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=NCCTC11637;  
RX MEDLINE=99255683; PubMed=10320570;  
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,  
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.,  
RT "Recombination and clonal groupings within Helicobacter pylori from  
RT different geographic regions.";  
Mol. Microbiol. 32:459-470(1999).  
DR EMBL; AJ239584; CAB37376.1;  
FT NON\_TER 1  
FT NON\_TER 148  
SQ SEQUENCE 148 AA; 16651 MW; 599C9C07C9B0DBIA CRC64;

Query Match 3.5%; Score 46; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 4.3e-39;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 NFTGGDLVDYNNQKATRLGQFNGNSFTSYKDSADTRTRVDNNAKNI 195  
Db 82 NFTGGDLVDYNNQKATRLGQFNGNSFTSYKDSADTRTRVDNNAKNI 127

RESULT 67  
Q9ZHU4 PRELIMINARY; PRT; 1291 AA.  
ID Q9ZHU4;  
AC Q9ZHU4;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=F78;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT nontoxic Helicobacter pylori.";  
J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049638; AAD04275.1;

SQ SEQUENCE 1291 AA; 139409 MW; 843C40D15DAA42DC CRC64;

Query Match 3.5%; Score 46; DB 2; Length 1291;  
Best Local Similarity 100.0%; Pred. No. 2.9e-38;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 992 IDSFARLQALKDQKFASLESAAEVLVYQFAPKYKPTNWNANAIGG 1037  
|||||  
DB 987 IDSFARLQALKDQKFASLESAAEVLVYQFAPKYKPTNWNANAIGG 1032

RESULT 68

ID Q9ZHV0 PRELIMINARY; PRT; 1293 AA.  
AC Q9ZHV0;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F44;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT noncytotoxic Helicobacter pylori.";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL: AF049627; AAD04266.1;  
SQ SEQUENCE 1293 AA; 139850 MW; F7876C59A9E93AF1 CRC64;

Query Match 3.5%; Score 46; DB 2; Length 1293;  
Best Local Similarity 100.0%; Pred. No. 2.9e-38;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 992 IDSFARLQALKDQKFASLESAAEVLVYQFAPKYKPTNWNANAIGG 1037  
|||||  
DB 987 IDSFARLQALKDQKFASLESAAEVLVYQFAPKYKPTNWNANAIGG 1032

RESULT 69

ID Q9ZHU0 PRELIMINARY; PRT; 1300 AA.  
AC Q9ZHU0;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F43;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT noncytotoxic Helicobacter pylori.";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL: AF049644; AAD04280.1;  
SQ SEQUENCE 1300 AA; 140394 MW; A6B91ECD980E0A1A CRC64;

Query Match 3.5%; Score 46; DB 2; Length 1300;  
Best Local Similarity 100.0%; Pred. No. 3e-38;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 992 IDSFARLQALKDQKFASLESAAEVLVYQFAPKYKPTNWNANAIGG 1037  
|||||  
DB 996 IDSFARLQALKDQKFASLESAAEVLVYQFAPKYKPTNWNANAIGG 1041

RESULT 70

Q9K2R6 PRELIMINARY; PRT; 139 AA.  
AC Q9K2R6;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=87A300, AND G104;  
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
RA Raviat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF050324; AAF26506.1;  
DR EMBL: AF050323; AAF26505.1;  
FT NON\_TER 1  
FT NON\_TER 139  
SQ SEQUENCE 139 AA; 14599 MW; B362B15353B7008B CRC64;

Query Match 3.5%; Score 45; DB 2; Length 139;  
Best Local Similarity 100.0%; Pred. No. 4.4e-38;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTDLWQSLNIIAPPEGYKDKPNTPSQSGAKNDKNESAKN 354  
|||||  
DB 40 IGTDLWQSLNIIAPPEGYKDKPNTPSQSGAKNDKNESAKN 84

RESULT 71

O32658 PRELIMINARY; PRT; 160 AA.  
ID O32658  
AC O32658;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HF006;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U63257; AAB61866.1;  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17936 MW; 783BAAB4D39C3853 CRC64;

Query Match 3.5%; Score 45; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 5e-38;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 VSGLLSWGLKQAEANKTPDKPKVWRIQAGKGFNEFPNKEYDLY 101  
 DB 1 VSGLLSWGLKQAEANKTPDKPKVWRIQAGKGFNEFPNKEYDLY 45

## RESULT 72

O32679 ID O32679 PRELIMINARY; PRT; 160 AA.  
 AC O32679;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HP093;  
 RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63278; AAB61887.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 160  
 SQ SEQUENCE 160 AA; 17888 MW; 0987EFA69FF90382 CRC64;

Query Match 3.5%; Score 45; DB 2; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 5e-38;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 VSGLLSWGLKQAEANKTPDKPKVWRIQAGKGFNEFPNKEYDLY 101  
 DB 1 VSGLLSWGLKQAEANKTPDKPKVWRIQAGKGFNEFPNKEYDLY 45

## RESULT 73

O32686 ID O32686 PRELIMINARY; PRT; 160 AA.  
 AC O32686;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HP168;  
 RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 13-160 FROM N.A.  
 RC STRAIN=CC46, CC25, CC30, CC33, AND CC35;  
 RX MEDLINE=98445420; PubMed=9770535;  
 RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,  
 RA Kunstmann E., Dyrek I., Achtman M.;  
 RT "Free recombination within Helicobacter pylori.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624 (1998).  
 RN [3]  
 RP SEQUENCE OF 13-160 FROM N.A.  
 RC STRAIN=5596;  
 RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,  
 RA Suerbaum S., Thompson S.A., van der Ende A., van Doorn L.J.;  
 RT "Recombination and clonal groupings within Helicobacter pylori from  
 different geographic regions";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63286; AAB61895.1; -;

DR EMBL; AJ009431; CAA08706.1; -;  
 DR EMBL; AJ239551; CAB37674.1; -;  
 DR EMBL; AJ009419; CAA08694.1; -;  
 DR EMBL; AJ009424; CAA08699.1; -;  
 DR EMBL; AJ009426; CAA08701.1; -;  
 DR EMBL; AJ009427; CAA08702.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 160  
 SQ SEQUENCE 160 AA; 17858 MW; E387E461807EFBEF CRC64;

Query Match 3.5%; Score 45; DB 2; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 5e-38;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 VSGLLSWGLKQAEANKTPDKPKVWRIQAGKGFNEFPNKEYDLY 101  
 DB 1 VSGLLSWGLKQAEANKTPDKPKVWRIQAGKGFNEFPNKEYDLY 45

## RESULT 74

O9R960 ID O9R960 PRELIMINARY; PRT; 380 AA.  
 AC O9R960;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=F67;  
 RX MEDLINE=98453456; PubMed=9780260;  
 RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
 RA Kuriyama M.;  
 RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
 RT noncytotoxic Helicobacter pylori.";  
 RL J. Infect. Dis. 178:1391-1398 (1998).  
 DR EMBL; AF049637; AAD04273.1; -;  
 FT NON\_TER 380  
 FT NON\_TER 380  
 SQ SEQUENCE 380 AA; 41047 MW; 136367A2327B0E0E CRC64;

Query Match 3.5%; Score 45; DB 2; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-37;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWMLQVYGLAPSYSTINTSKVTGEVNFHNLTVGD 294  
 DB 250 SVKLMGNVWMLQVYGLAPSYSTINTSKVTGEVNFHNLTVGD 294

## RESULT 75

O9F9G4 ID O9F9G4 PRELIMINARY; PRT; 826 AA.  
 AC O9F9G4;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=61;  
 RA Kuo C.H., Wang W.C.;



RT \*Genetic and functional comparison of Helicobacter pylori vacuolating  
RT toxin between sla/mlr and sla/m2 isolates.\*;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF195011; AAG28439.1; -  
FT NON\_TER 1  
FT NON\_TER 826 826  
SQ SEQUENCE 826 AA; 88669 MW; B166842EC5A34DFC CRC64;

Query Match 3.5%; Score 45; DB 2; Length 826;  
Best Local Similarity 100.0%; Pred. No. 2.2e-37;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWVGRLQVGVAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||  
DB 217 SVKLMGNVWVGRLQVGVAYLAPSYSTINTSKVTGEVNFNHLTVGD 261  
|||||

RESULT 76  
QYF9G0 PRELIMINARY; PRT; 829 AA.  
ID QYF9G0  
AC QYF9G0  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=99;  
RA Kuo C.H., Wang W.C.;  
RT \*Genetic and functional comparison of Helicobacter pylori vacuolating  
RT toxin between sla/mlr and sla/m2 isolates.\*;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF195015; AAG28443.1; -  
FT NON\_TER 1  
FT NON\_TER 829 829  
SQ SEQUENCE 829 AA; 89106 MW; 2DB64610DCE207DE CRC64;

Query Match 3.5%; Score 45; DB 2; Length 829;  
Best Local Similarity 100.0%; Pred. No. 2.2e-37;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWVGRLQVGVAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||  
DB 217 SVKLMGNVWVGRLQVGVAYLAPSYSTINTSKVTGEVNFNHLTVGD 261  
|||||

RESULT 77  
QYF9F8 PRELIMINARY; PRT; 829 AA.  
ID QYF9F8  
AC QYF9F8  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=184;  
RA Kuo C.H., Wang W.C.;  
RT \*Genetic and functional comparison of Helicobacter pylori vacuolating  
RT toxin between sla/mlr and sla/m2 isolates.\*;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF195017; AAG28445.1; -  
FT NON\_TER 1  
FT NON\_TER 829 829  
SQ SEQUENCE 829 AA; 89339 MW; BDE651E23A57EECE CRC64;

Query Match 3.5%; Score 45; DB 2; Length 829;  
Best Local Similarity 100.0%; Pred. No. 2.2e-37;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWVGRLQVGVAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||  
DB 217 SVKLMGNVWVGRLQVGVAYLAPSYSTINTSKVTGEVNFNHLTVGD 261  
|||||

RESULT 78  
QYF9G5 PRELIMINARY; PRT; 861 AA.  
ID QYF9G5  
AC QYF9G5  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=55;  
RA Kuo C.H., Wang W.C.;  
RT \*Genetic and functional comparison of Helicobacter pylori vacuolating  
RT toxin between sla/mlr and sla/m2 isolates.\*;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF195010; AAG28438.1; -  
FT NON\_TER 1  
FT NON\_TER 861 861  
SQ SEQUENCE 861 AA; 92671 MW; C0C684E752E930C5 CRC64;

Query Match 3.5%; Score 45; DB 2; Length 861;  
Best Local Similarity 100.0%; Pred. No. 2.2e-37;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWVGRLQVGVAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||  
DB 217 SVKLMGNVWVGRLQVGVAYLAPSYSTINTSKVTGEVNFNHLTVGD 261  
|||||

RESULT 79  
QYF9F7 PRELIMINARY; PRT; 861 AA.  
ID QYF9F7  
AC QYF9F7  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1939;  
RA Kuo C.H., Wang W.C.;  
RT \*Genetic and functional comparison of Helicobacter pylori vacuolating  
RT toxin between sla/mlr and sla/m2 isolates.\*;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF195018; AAG28446.1; -  
FT NON\_TER 1  
FT NON\_TER 861 861

SO SEQUENCE 861 AA; 92852 MW; 8F07C1710DC20E82 CRC64;

Query Match 3.5%; Score 45; DB 2; Length 861;  
Best Local Similarity 100.0%; Pred. No. 2.2e-37;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRQYVGYAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||  
DB 217 SVKLMGNVWGRQYVGYAYLAPSYSTINTSKVTGEVNFNHLTVGD 261  
|||||

RESULT 80

ID Q9F9G1 PRELIMINARY; PRT; 864 AA.  
AC Q9F9G1;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=92;  
RA Kuo C.H., Wang W.C.;  
RT "Genetic and functional comparison of Helicobacter pylori vacuolating toxin between sla/mlt and sla/m2 isolates.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF195014; AAG28442.1;  
FT NON\_TER 1  
FT NON\_TER 864  
SQ SEQUENCE 864 AA; 93226 MW; 2CE3D82328E7CA35 CRC64;

Query Match 3.5%; Score 45; DB 2; Length 864;  
Best Local Similarity 100.0%; Pred. No. 2.3e-37;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRQYVGYAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||  
DB 217 SVKLMGNVWGRQYVGYAYLAPSYSTINTSKVTGEVNFNHLTVGD 261  
|||||

RESULT 81

ID Q9F9G6 PRELIMINARY; PRT; 866 AA.  
AC Q9F9G6;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=50;  
RA Kuo C.H., Wang W.C.;  
RT "Genetic and functional comparison of Helicobacter pylori vacuolating toxin between sla/mlt and sla/m2 isolates.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF195009; AAG28437.1;  
FT NON\_TER 1  
FT NON\_TER 866  
SQ SEQUENCE 866 AA; 93254 MW; 86B30EFA3D73619 CRC64;

Query Match 3.5%; Score 45; DB 2; Length 866;  
Best Local Similarity 100.0%; Pred. No. 2.3e-37;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRQYVGYAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||  
DB 217 SVKLMGNVWGRQYVGYAYLAPSYSTINTSKVTGEVNFNHLTVGD 261  
|||||

RESULT 82

ID Q9R964 PRELIMINARY; PRT; 1082 AA.  
AC Q9R964;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F31;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y., Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and noncytotoxic Helicobacter pylori.";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049623; AAD04262.1;  
FT NON\_TER 1082  
FT NON\_TER 1082  
SQ SEQUENCE 1082 AA; 116324 MW; 0C1545742B8AAF5B CRC64;

Query Match 3.5%; Score 45; DB 2; Length 1082;  
Best Local Similarity 100.0%; Pred. No. 2.8e-37;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRQYVGYAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||  
DB 250 SVKLMGNVWGRQYVGYAYLAPSYSTINTSKVTGEVNFNHLTVGD 294  
|||||

RESULT 83

ID Q9ZHV3 PRELIMINARY; PRT; 1291 AA.  
AC Q9ZHV3;  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F21;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y., Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and noncytotoxic Helicobacter pylori.";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049620; AAD04261.1;  
FT NON\_TER 1  
FT NON\_TER 1291  
SQ SEQUENCE 1291 AA; 139539 MW; A84CEADF43B6D4D1 CRC64;

Query Match 3.5%; Score 45; DB 2; Length 1291;  
Best Local Similarity 100.0%; Pred. No. 3.2e-37;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

Db 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

# RESULT 84

Q92HU9 ID Q92HU9 PRELIMINARY; PRT; 1291 AA.

AC Q92HU9

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)

DE VACUOLATING CYTOTOXIN PRECURSOR.

GN VACA.

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

OC Helicobacter.

OX NCBI\_TaxID=210;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=F45;

RX MEDLINE=98453456; PubMed=9780260;

RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,

RA Kuriyama M.;

RT "Full-length sequence analysis of the vacA gene from cytotoxic and

RT nontoxic Helicobacter pylori.";

RL J. Infect. Dis. 178:1391-1398(1998).

DR EMBL; AF049628; AAD04267.1; -

SQ SEQUENCE 1291 AA; 139407 MW; B560563C15BD77A2 CRC64;

## Query Match

Best Local Similarity 100.0%; Pred. No. 3.2e-37; Length 1291;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

Db 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

# RESULT 85

Q92HT9

ID Q92HT9 PRELIMINARY; PRT; 1291 AA.

AC Q92HT9

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)

DE VACUOLATING CYTOTOXIN PRECURSOR.

GN VACA.

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

OC Helicobacter.

OX NCBI\_TaxID=210;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=F61;

RX MEDLINE=98453456; PubMed=9780260;

RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,

RA Kuriyama M.;

RT "Full-length sequence analysis of the vacA gene from cytotoxic and

RT nontoxic Helicobacter pylori.";

RL J. Infect. Dis. 178:1391-1398(1998).

DR EMBL; AF049645; AAD04281.1; -

SQ SEQUENCE 1291 AA; 139441 MW; A5977DF3453EB4D5 CRC64;

## Query Match

Best Local Similarity 100.0%; Pred. No. 3.2e-37; Length 1291;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

Db 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

Db 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

# RESULT 86

Q92HT5

ID Q92HT5 PRELIMINARY; PRT; 1291 AA.

AC Q92HT5

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)

DE VACUOLATING CYTOTOXIN PRECURSOR.

GN VACA.

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

OC Helicobacter.

OX NCBI\_TaxID=210;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=F69;

RX MEDLINE=98453456; PubMed=9780260;

RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,

RA Kuriyama M.;

RT "Full-length sequence analysis of the vacA gene from cytotoxic and

RT nontoxic Helicobacter pylori.";

RL J. Infect. Dis. 178:1391-1398(1998).

DR EMBL; AF049649; AAD04285.1; -

SQ SEQUENCE 1291 AA; 139363 MW; C568C500A2EEB8CF CRC64;

## Query Match

Best Local Similarity 100.0%; Pred. No. 3.2e-37; Length 1291;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

Db 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

# RESULT 87

Q9R962

ID Q9R962 PRELIMINARY; PRT; 1291 AA.

AC Q9R962

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)

DE VACUOLATING CYTOTOXIN PRECURSOR.

GN VACA.

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

OC Helicobacter.

OX NCBI\_TaxID=210;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=F47;

RX MEDLINE=98453456; PubMed=9780260;

RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,

RA Kuriyama M.;

RT "Full-length sequence analysis of the vacA gene from cytotoxic and

RT nontoxic Helicobacter pylori.";

RL J. Infect. Dis. 178:1391-1398(1998).

DR EMBL; AF049629; AAD04288.1; -

SQ SEQUENCE 1291 AA; 139347 MW; E54C91429E9DF920 CRC64;

## Query Match

Best Local Similarity 100.0%; Pred. No. 3.2e-37; Length 1291;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

Db 250 SVKLMGNVWGRLOQYVGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294

AC	Q9ZHT7;
DT	01-MAY-1999 (TrEMBLrel. 10, Created)
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE	VACUOLATING CYTOTOXIN PRECURSOR.
GN	VACA.
OS	Helicobacter pylori (Campylobacter pylori).
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC	Helicobacter.
OX	NCBI_TaxID=210;
[1]	SEQUENCE FROM N.A..
RN	STRAIN-F68;
RC	MEDLINE=98453456; PubMed=9780260;
RA	Ito Y., Azuma T., Ito S., Suto H., Miyaaji H., Yamazaki Y., Kohli Y.,
RA	Kuriyama M.;
RT	"Full-length sequence analysis of the vacA gene from cytotoxic and
RT	nontoxic Helicobacter pylori.";
RL	J. Infect. Dis. 178:1391-1398(1998).
DR	EMBL; AF049648; AAD04283.1; -
SQ	SEQUENCE 1296 AA; 140049 MW; E66421537383F709 CRC64;
Query Match            3.5%; Score 45; DB 2; Length 1296;	
Best Local Similarity 100.0%; Pred. No. 3.2e-37;	
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps	
QY	250 SVKLMGNWNGRLOYGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294 
Db	250 SVKLMGNWNGRLOYGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294 
RESULT 91:	
Q9KJAG	PRELIMINARY; PRT; 1303 AA.
ID	Q9KJAG
AC	Q9KJAG;
DC	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
GN	VACA.
OS	Helicobacter pylori (Campylobacter pylori).
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC	Helicobacter.
OX	NCBI_TaxID=210;
[1]	SEQUENCE FROM N.A.
RN	STRAIN=CH2;
RC	MEDLINE=20278176; PubMed=10816542;
RA	Ji X., Fernandez T., Burroni D., Pagliaccia C., Atherton J.C.,
RA	Reyrat J.M., Rappuoli R., Telford J.L.;
RT	"Cell specificity of Helicobacter pylori cytotoxin is determined by a
RT	short region in the polymorphic midregion.";
RL	Infect. Immun. 68:3754-3757(2000).
DR	EMBL; AF191639; AAF86763.1; -
DR	InterPro: IPR001029; -
DR	ProDom: PD000316; -; 1.
SQ	SEQUENCE 1303 AA; 141069 MW; ECFDF290125D878D CRC64;
Query Match            3.5%; Score 45; DB 2; Length 1303;	
Best Local Similarity 100.0%; Pred. No. 3.3e-37;	
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps	
QY	250 SVKLMGNWNGRLOYGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294 
Db	250 SVKLMGNWNGRLOYGAYLAPSYSTINTSKVTGEVNFNHLTVGD 294 
RESULT 92	
O87749	PRELIMINARY; PRT; 148 AA.
ID	O87749
AC	O87749;

DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CC10;  
 RX MEDLINE=98445420; PubMed=9770535;  
 RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,  
 RA Kunstmann E., Dyrek I., Achtman M.;  
 RT "Free recombination within Helicobacter pylori.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).  
 DR EMBL; AJ009417; CAA08692.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 148  
 SQ SEQUENCE 148 AA; 16587 MW; A3F77FED3795D81F CRC64;

Query Match 3.4%; Score 44; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-37;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
 |||||  
 DB 77 LSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 120

RESULT 93  
 Q9R853  
 ID Q9R853 PRELIMINARY; PRT; 148 AA.  
 AC Q9R853;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J166;  
 RX MEDLINE=99255683; PubMed=10320570;  
 RA Achermann M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,  
 RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;  
 RT "Recombination and clonal groupings within Helicobacter pylori from  
 RT different geographic regions.";   
 RL Mol. Microbiol. 32:459-470(1999).  
 DR EMBL; AJ239552; CAB37675.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 148  
 SQ SEQUENCE 148 AA; 16644 MW; 5EFF36034690C429 CRC64;

Query Match 3.4%; Score 44; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-37;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
 |||||  
 DB 77 LSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 120

RESULT 94  
 Q9R853  
 ID Q9R853 PRELIMINARY; PRT; 148 AA.  
 AC Q9R853;

DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CC26;  
 RX MEDLINE=98445420; PubMed=9770535;  
 RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,  
 RA Kunstmann E., Dyrek I., Achtman M.;  
 RT "Free recombination within Helicobacter pylori.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).  
 DR EMBL; AJ009420; CAA08695.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 148  
 SQ SEQUENCE 148 AA; 16631 MW; 2692247F3F760368 CRC64;

Query Match 3.4%; Score 44; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-37;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
 |||||  
 DB 77 LSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 120

RESULT 95  
 Q9R843  
 ID Q9R843 PRELIMINARY; PRT; 148 AA.  
 AC Q9R843;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CC8;  
 RX MEDLINE=98445420; PubMed=9770535;  
 RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,  
 RA Kunstmann E., Dyrek I., Achtman M.;  
 RT "Free recombination within Helicobacter pylori.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).  
 DR EMBL; AJ009438; CAA08713.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 148  
 SQ SEQUENCE 148 AA; 16589 MW; AAA6023D5B505744 CRC64;

Query Match 3.4%; Score 44; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-37;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
 |||||  
 DB 77 LSGLRNFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 120

RESULT 96  
 Q9R3L6  
 ID Q9R3L6 PRELIMINARY; PRT; 148 AA.  
 AC Q9R3L6;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC28;  
RX MEDLINE=99255683; PubMed=10320570;  
RA Achman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,  
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.,  
RT "Recombination and clonal groupings within Helicobacter pylori from  
RT different geographic regions."  
RL Mol. Microbiol. 32:459-470(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC28, AND CC3;  
RX MEDLINE=98445420; PubMed=9770535;  
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,  
RA Kunstmann E., Dyrek I., Achtman M.,  
RT "Free recombination within Helicobacter pylori."  
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).  
DR EMBL: AJ239553; CAB37676.1; -  
DR EMBL: AJ009421; CAA08696.1; -  
DR EMBL: AJ009429; CAA08704.1; -  
FT NON\_TER 1  
FT NON\_TER 148  
SQ SEQUENCE 148 AA; 16630 MW; 5EFF37D83690C429 CRC64;

Query Match 3.4%; Score 44; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 5.2e-37;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
Db 77 LSGLRFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 120

RESULT 97  
Q32657  
ID 032657 PRELIMINARY; PRT; 160 AA.  
AC 032657;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP003;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U63256; AAB61865.1; -  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17886 MW; 018AF9968DC837B2 CRC64;

Query Match 3.4%; Score 44; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 5.5e-37;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
Db 89 LSGLRFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 132

RESULT 98  
Q32673  
ID 032673 PRELIMINARY; PRT; 160 AA.  
AC 032673;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP059;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U63272; AAB61881.1; -  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17941 MW; 24ESB3D98FD64E18 CRC64;

Query Match 3.4%; Score 44; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 5.5e-37;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 TGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRVDFNAKNI 195  
Db 96 TGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRVDFNAKNI 139

RESULT 99  
Q32675  
ID 032675 PRELIMINARY; PRT; 160 AA.  
AC 032675;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP061;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U63274; AAB61883.1; -  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17900 MW; 018AF94DFDC837B2 CRC64;

Query Match 3.4%; Score 44; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 5.5e-37;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
Db 89 LSGLRFTGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 132

RESULT 100  
Q9RE74  
ID Q9RE74 PRELIMINARY; PRT; 96 AA.  
AC Q9RE74;  
DT 01-MAY-2000 (Tremblrel. 13, Created)

FT	NON_TER	1	1	
FT	NON_TER	130	130	
SQ	SEQUENCE	130 AA:	13460 MW:	02BABA98F222D5C1 CRC64:

Query Match 3.3%; Score 43; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 5e-36;  
Matches 43; Conservative 0; Mismatches 0; Indels

QY	485	DTKNGTATFNNDISLGRFVNLKYDAHTANFKCIDTGNCGGNTL	527
Db	88	DTKNGTATFNNDISLGRFVNLKYDAHTANFKCIDTGNCGGNTL	130

RESULT 104  
091.808

ID	Q9LB08	PRELIMINARY;	PRT;	78 AA.
AC	Q9LB08;			
DT	01-OCT-2000	(TrEMBLrel. 15, Created)		
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)		
DT	01-OCT-2000	(TrEMBLrel. 15, Last annotation update)		
DE	VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).			
GN	VACA.			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;			
OC	Helicobacter			
OX	NCBI_TaxID-210;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-INDIA29;			
RC	MEDLINE=20270153; PubMed=10809703;			
RA	Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,			
RA	Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.			
RA	Nair G.B., Berg D.E.;			
RT	"Distinctiveness of genotypes of Helicobacter pylori in Calcutta,			
RT	India."			
RL	J. Bacteriol. 182:3219-3227(2000).			
DR	EMBL: AF217734; AAF42845.1; -.			
FT	NON_TER 1 1			
FT	NON_TER 78 78			
FT	SEQUENCE 78 AA; 8044 MW; 5A2C8092DC5E5305 CRC64:			
SQ				

Query Match 3.2%; Score 42; DB 2; Length 78;  
Best Local Similarity 100.0%; Pred. No. 3.5e-35;  
Matches 42; Conservative 0; Mismatches 0; Indels

10 INRPLVSLALVGALVSTPQQSHAAFTTIIIPAIVGSIATG 51  
11 |||||  
12 |||||  
13 |||||  
14 |||||  
15 |||||  
16 |||||  
17 |||||  
18 |||||  
19 |||||  
20 |||||  
21 |||||  
22 |||||  
23 |||||  
24 |||||  
25 |||||  
26 |||||  
27 |||||  
28 |||||  
29 |||||  
30 |||||  
31 |||||  
32 |||||  
33 |||||  
34 |||||  
35 |||||  
36 |||||  
37 |||||  
38 |||||  
39 |||||  
40 |||||  
41 |||||  
42 |||||  
43 |||||  
44 |||||  
45 |||||  
46 |||||  
47 |||||  
48 |||||  
49 |||||  
50 |||||  
51 |||||  
52 |||||  
53 |||||  
54 |||||  
55 |||||  
56 |||||  
57 |||||  
58 |||||  
59 |||||  
60 |||||  
61 |||||  
62 |||||  
63 |||||  
64 |||||  
65 |||||  
66 |||||  
67 |||||  
68 |||||  
69 |||||  
70 |||||  
71 |||||  
72 |||||  
73 |||||  
74 |||||  
75 |||||  
76 |||||  
77 |||||  
78 |||||  
79 |||||  
80 |||||  
81 |||||  
82 |||||  
83 |||||  
84 |||||  
85 |||||  
86 |||||  
87 |||||  
88 |||||  
89 |||||  
90 |||||  
91 |||||  
92 |||||  
93 |||||  
94 |||||  
95 |||||  
96 |||||  
97 |||||  
98 |||||  
99 |||||  
100 |||||

RESULT 105  
086253

ID	086253	PRELIMINARY;	PRT;	403 AA.
AC	086253;			
DT	01-NOV-1998	(T=EMBLrel. 08,	Created)	
DD	01-NOV-1998	(T=EMBLrel. 08,	Last sequence update)	
DT	01-JUN-2000	(T=EMBLrel. 14,	Last annotation update)	
DE	VACA (FRAGMENT).			
DE	VACA.			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;			
OC	Helicobacter.			
OC	NCBI_TaxID=210;			
NRN	[1]			
RPN	SEQUENCE FROM N.A.			
RC	STRAIN-M229A;			
RRX	MEDLINE=98180430; PubMed=9521135;			
RRX	Han S.R., Schreiber H.J., Bhakdi S., Loos M., Maeurer M.J.;			
RRX	"vaca genotypes and genetic diversity in clinical isolates of			
RRX	Helicobacter pylori.";			
RL	Clin. Diagn. Lab. Immunol. 5:139-145(1998).			
RT	EMBL: AJ006970; CAA07359.1; -.			
TR	NON_TER. 1			
TR	NON_TER. 1			
TR	NON_TER. 403			
TR	NON_TER. 403			
SO	SEQUENCE 403 AA, 43728 MW:			
SO	3C2ERABI59AR225B CRC64			

```
Query Match      3.2%; Score 42; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.5e-34;
Matches 42; Conservative 0; Mismatches 0; Indels
```

QY 447 VENLTGNTVTDGRLRVNNGGYALAGSSANFEKAGTDTKN 488  
 |||||  
 Db 23 VENLTGNTVTDGRLRVNNGGYALAGSSANFEKAGTDTKN 64

RESULT 106

Q9R850	PRELIMINARY;	PRT;	148 AA.
AC	Q9R850		
ID	Q9R850;		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2000	(TREMBLrel. 14, Last annotation update)	
DE	VACUOLATING CYTOTOXIN (FRAGMENT).		
GN	VACA.		
OS	Helicobacter pylori (Campylobacter pylori).		
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;		
OC	Helicobacter.		
OX	NCBI_TaxID=210;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CC31;		
RX	MEDLINE=98445420; PubMed=9770535;		
RA	Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,		
RT	Kunstmann E., Dyrek I., Achtman M.;		
RT	"Free recombination within Helicobacter pylori."		
RL	Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).		
DR	EMBL; AJ009425; CA008700.1; -.		
FT	NON_TER	1	
FT	NON_TER	148	
SQ	SEQUENCE	148 AA;	16543 MW; 1AFA4AEA9E246370 CRC64;

Query Match 3.2%; Score 41; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 6.8e-34;  
Matches 41; Conservative 0; Mismatches 0; Indels

Qy	103	SLSSKIDGGWDGNAARHYVWGGQNKLEVDMKDAVCTY	143
Db	35	SLSSKIDGGWDGNAARHYVWGGQNKLEVDMKDAVCTY	75

RESULT 107  
032665

032665	PRELIMINARY;	PRT;	160 AA.
032665;			
01-JAN-1998	(TReMBLrel. 05, Created)		
01-JAN-1998	(TReMBLrel. 05, Last sequence update)		
01-JUN-2000	(TReMBLrel. 14, Last annotation update)		
DE	VACUOLATING CYTOTOXIN (FRAGMENT).		
VACA.			
GN	Helicobacter pylori (Campylobacter pylori).		
OS	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;		
OC	Helicobacter.		
OC	Helicobacter.		
OX	NCBI_TaxID=210;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HF025;		
RC	Besch R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;		
DR	Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; U63264; AB61873.1; -.		
FT	NON_TER	1	
FT	NON_TER	160	
SQ	SEQUENCE	160 AA;	B39249BCBB27F486 CRC64;

Query Match 3.2%; Score 41; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 7.3e-34;  
Matches 41; Conservative 0; Mismatches 0; Indels



QY 103 SLSSKIDGWDGNGNAARHYWVGQGNKLEVDMDKDAVGTY 143  
 |||||  
 Db 47 SLSSKIDGWDGNGNAARHYWVGQGNKLEVDMDKDAVGTY 87

RESULT 108  
 O32667 PRELIMINARY; PRT; 160 AA.  
 ID O32667  
 AC O32667;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HP037;  
 RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63266; AAB61875.1;  
 FT NON\_TER 1  
 FT NON\_TER 160  
 SQ SEQUENCE 160 AA; 17798 MW; 17EE86B019F2817F CRC64;

Query Match 3.2%; Score 41; DB 2; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-34;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 SLSSKIDGWDGNGNAARHYWVGQGNKLEVDMDKDAVGTY 143  
 |||||  
 Db 47 SLSSKIDGWDGNGNAARHYWVGQGNKLEVDMDKDAVGTY 87

## RESULT 109

ID O85022 PRELIMINARY; PRT; 240 AA.  
 AC O85022;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R10A;  
 RX MEDLINE=98314552; PubMed-9652444;  
 RA Pan Z.J., Berg D.E., van der Hulst R.W., Su W.W., Raudoniklene A.,  
 Xiao S.D., Dankert J., Tytgat G.N., van der Ende A.;  
 RT "Prevalence of vacuolating cytotoxin production and distribution of  
 distinct vacA alleles in Helicobacter pylori from China."  
 RL J. Infect. Dis. 178:220-226(1998).  
 DR EMBL; AF035609; AAC31125.1;  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SQ SEQUENCE 240 AA; 26722 MW; 9402FF0DDE0A235A CRC64;

Query Match 3.2%; Score 41; DB 2; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 1e-33;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNWYFDARNIKNVEITNKLAFGQSPWGTSKLMFNNTLG 657  
 |||||

Db 77 PNWYFDARNIKNVEITNKLAFGQSPWGTSKLMFNNTLG 117

## RESULT 110

O85029 PRELIMINARY; PRT; 240 AA.  
 ID O85029  
 AC O85029;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R34A;  
 RX MEDLINE=98314552; PubMed-9652444;  
 RA Pan Z.J., Berg D.E., van der Hulst R.W., Su W.W., Raudoniklene A.,  
 Xiao S.D., Dankert J., Tytgat G.N., van der Ende A.;  
 RT "Prevalence of vacuolating cytotoxin production and distribution of  
 distinct vacA alleles in Helicobacter pylori from China."  
 RL J. Infect. Dis. 178:220-226(1998).  
 DR EMBL; AF035616; AAC31132.1;  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SQ SEQUENCE 240 AA; 26765 MW; 7626877CC220B8C5 CRC64;

Query Match 3.2%; Score 41; DB 2; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 1e-33;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNWYFDARNIKNVEITNKLAFGQSPWGTSKLMFNNTLG 657  
 |||||  
 Db 77 PNWYFDARNIKNVEITNKLAFGQSPWGTSKLMFNNTLG 117

## RESULT 111

ID O9X408 PRELIMINARY; PRT; 240 AA.  
 AC O9X408;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=T-34;  
 RX Lin C.W.;  
 RA "Helicobacter pylori Taiwanese isolates."  
 RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF091815; AAD3352.1;  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SQ SEQUENCE 240 AA; 26651 MW; 2FABB3F1C0E74946 CRC64;

Query Match 3.2%; Score 41; DB 2; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 1e-33;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNWYFDARNIKNVEITNKLAFGQSPWGTSKLMFNNTLG 657  
 |||||  
 Db 77 PNWYFDARNIKNVEITNKLAFGQSPWGTSKLMFNNTLG 117

## RESULT 112

O34109 ID O34109 PRELIMINARY; PRT; 244 AA.  
 AC O34109;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE STRAIN F94 VACA (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=F94;  
 RX MEDLINE=97339580; PubMed=9196179;  
 RA Ito Y., Azuma T., Ito S., Miyaji H., Hirai M., Yamazaki Y., Sato F.,  
 RA Kato T., Kohli Y., Kuriyama M.;  
 RT "Analysis and typing of the vacA gene from cagA-positive strains of  
 RT Helicobacter pylori isolated in Japan.";  
 RL J. Clin. Microbiol. 35:1710-1714(1997).  
 DR EMBL; U91577; AAB66699.1; -;  
 FT NON\_TER 1  
 FT 244  
 FT NON\_TER 244  
 SQ SEQUENCE 244 AA; 26599 MW; 13275CDBE9432E6B CRC64;

Query Match 3.2%; Score 41; DB 2; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-33;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 617 PNWYDARNIKNVEITNKLAFGPQSPWGTSKLMFNNTLG 657  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 101 PNWYDARNIKNVEITNKLAFGPQSPWGTSKLMFNNTLG 141

## RESULT 113

O87487 ID O87487 PRELIMINARY; PRT; 406 AA.  
 AC O87487;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING TOXIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=549;  
 RX MEDLINE=98440809; PubMed=9753641;  
 RA Wang H.J., Chang P.C.L., Kuo C.H., Tzeng C.S., Wang W.C.;  
 RT "Characterization of the C-terminal domain of Helicobacter pylori  
 RT vacuolating toxin and its relationship with extracellular toxin  
 RT production.";  
 RL Biochem. Biophys. Res. Commun. 250:397-402(1998).  
 DR EMBL; AF077941; AAC63523.1; -;  
 FT NON\_TER 1  
 FT 406  
 FT NON\_TER 406  
 SQ SEQUENCE 406 AA; 44186 MW; 4FF6002A95B3B5E7 CRC64;

Query Match 3.2%; Score 41; DB 2; Length 406;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-33;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 891 SVFELANSKIDITLYANSAGQGRDLLQTLIDSHDAGYAR 931  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 1 SVFELANSKIDITLYANSAGQGRDLLQTLIDSHDAGYAR 41

## RESULT 114

Q92HU1 ID Q92HU1 PRELIMINARY; PRT; 1291 AA.  
 AC Q92HU1;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR.  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=F38;  
 RX MEDLINE=98453456; PubMed=9780260;  
 RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
 RA Kuriyama M.;  
 RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
 RT noncytotoxic Helicobacter pylori.";  
 RL J. Infect. Dis. 178:1391-1398(1998).  
 DR EMBL; AF049643; RAD04278.1; -;  
 SQ SEQUENCE 1291 AA; 139427 MW; 08091C7A0C46EC16 CRC64;

Query Match 3.2%; Score 41; DB 2; Length 1291;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-33;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 617 PNWYDARNIKNVEITNKLAFGPQSPWGTSKLMFNNTLG 657  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 612 PNWYDARNIKNVEITNKLAFGPQSPWGTSKLMFNNTLG 652

## RESULT 115

Q92HU3 ID Q92HU3 PRELIMINARY; PRT; 1296 AA.  
 AC Q92HU3;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR.  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=F33;  
 RX MEDLINE=98453456; PubMed=9780260;  
 RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
 RA Kuriyama M.;  
 RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
 RT noncytotoxic Helicobacter pylori.";  
 RL J. Infect. Dis. 178:1391-1398(1998).  
 DR EMBL; AF049641; RAD04276.1; -;  
 SQ SEQUENCE 1296 AA; 139963 MW; A47D7FC471925596 CRC64;

Query Match 3.2%; Score 41; DB 2; Length 1296;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-33;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 617 PNWYDARNIKNVEITNKLAFGPQSPWGTSKLMFNNTLG 657  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 617 PNWYDARNIKNVEITNKLAFGPQSPWGTSKLMFNNTLG 657

## RESULT 116

O34112

ID O34112 PRELIMINARY; PRT; 244 AA.  
AC O34112;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE STRAIN F80 VACA (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F80;  
RX MEDLINE=97339580; PubMed=9196179;  
RA Ito Y., Azuma T., Ito S., Miyaji H., Hirai M., Yamazaki Y., Sato F.,  
RA Kato Y., Kohli Y., Kuriyama M.;  
RT "Analysis and typing of the vacA gene from cagA-positive strains of  
RT Helicobacter pylori isolated in Japan.";  
RL J. Clin. Microbiol. 35:1710-1714(1997).  
DR EMBL; U91580; AAB66700.1; -;  
FT NON\_TER 1 1  
FT NON\_TER 244 244  
SQ SEQUENCE 244 AA; 26429 MW; 19953D0A0C4DD3E8 CRC64;

Query Match 3.1%; Score 40; DB 2; Length 244;  
Best Local Similarity 100.0%; Pred. No. 1.2e-32;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 VGNAAAMFNNDISATGYKPLIKINSQDLIKNTEHYL 736  
DB 181 VGNAAAMFNNDISATGYKPLIKINSQDLIKNTEHYL 220

RESULT 117  
O87484 PRELIMINARY; PRT; 406 AA.  
ID O87484;  
AC O87484;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING TOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=50;  
RX MEDLINE=98440809; PubMed=9753641;  
RA Wang H.J., Chang P.C.L., Kuo C.H., Tzeng C.S., Wang W.C.;  
RT "Characterization of the C-terminal domain of Helicobacter pylori  
RT vacuolating toxin and its relationship with extracellular toxin  
RT production.";  
RL Biochem. Biophys. Res. Commun. 250:397-402(1998).  
DR EMBL; AF077938; AAC63520.1; -;  
FT NON\_TER 1 1  
SQ SEQUENCE 406 AA; 44286 MW; 5D16899BAC5BC05E CRC64;

Query Match 3.1%; Score 40; DB 2; Length 406;  
Best Local Similarity 100.0%; Pred. No. 1.8e-32;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 998 RLQALKDKQFASLESAAEVLQFAPKYKPTNWNNAIGG 1037  
DB 108 RLQALKDKQFASLESAAEVLQFAPKYKPTNWNNAIGG 147

RESULT 118  
Q92HU6

ID Q92HU6 PRELIMINARY; PRT; 1291 AA.  
AC Q92HU6;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F56;  
RX MEDLINE=98453456; PubMed=9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT noncytotoxic Helicobacter pylori.";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049633; AAD04271.1; -;  
SQ SEQUENCE 1291 AA; 139439 MW; FD05AA63664A7158 CRC64;

Query Match 3.1%; Score 40; DB 2; Length 1291;  
Best Local Similarity 100.0%; Pred. No. 5.1e-32;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 QOTHRKNRPVSLVGLVGVLSITPQOSHAAFTTVIIPA 43  
DB 4 QOTHRKNRPVSLVGLVGVLSITPQOSHAAFTTVIIPA 43

RESULT 119  
O32662 PRELIMINARY; PRT; 160 AA.  
ID O32662;  
AC O32662;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HF020;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63261; AAB61870.1; -;  
FT NON\_TER 1 1  
FT NON\_TER 160 160  
SQ SEQUENCE 160 AA; 17641 MW; A94FD14F2BDC49AC CRC64;

Query Match 3.0%; Score 39; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 8.8e-32;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 WGLQKQAEANKTPDKPVWRIQAGKGFNFPNKEYDLY 101  
DB 7 WGLQKQAEANKTPDKPVWRIQAGKGFNFPNKEYDLY 45

RESULT 120  
O32682 PRELIMINARY; PRT; 160 AA.  
ID O32682;  
AC O32682;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE VACUOLATING CYTOTOXIN (FRAGMENT);\*  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP131;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63281; AAB61890.1; -  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17873 MW; 52199578B0743BAF CRC64;

Query Match 3.0%; Score 39; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 8.e-32;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 AVGTYTLGLRNFSGDLDVNMOKATRLGQFNGNSFTS 177  
|||||  
DB 83 AVGTYTLGLRNFSGDLDVNMOKATRLGQFNGNSFTS 121  
|||||

RESULT 121  
Q9S314  
ID Q9S314 PRELIMINARY; PRT; 148 AA.  
AC Q9S314;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5-1;  
RX MEDLINE=99255683; PubMed=10320570;  
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,  
Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;  
RT "Recombination and clonal groupings within Helicobacter pylori from  
different geographic regions."  
RL Mol. Microbiol. 32:459-470(1999).  
DR EMBL; AJ239554; CAB37677.1; -  
FT NON\_TER 1  
FT NON\_TER 148  
SQ SEQUENCE 148 AA; 16503 MW; 53BAB7E35823496B CRC64;

Query Match 2.9%; Score 38; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 9.e-31;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFSGDLDVNMOKATRLGQFNGNSFTSYKDSA 182  
|||||  
DB 77 LSGLRNFSGDLDVNMOKATRLGQFNGNSFTSYKDSA 114  
|||||

RESULT 122  
Q9R849  
ID Q9R849 PRELIMINARY; PRT; 148 AA.  
AC Q9R849;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC36;  
RX MEDLINE=98445420; PubMed=9770535;  
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,  
Kunstmann E., Dyrek I., Achtman M.;  
RL "Free recombination within Helicobacter pylori."  
RT Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).  
DR EMBL; AJ009428; CAA08703.1; -  
FT NON\_TER 1  
FT NON\_TER 148  
SQ SEQUENCE 148 AA; 16602 MW; 9E6124759FDC10CC CRC64;

Query Match 2.9%; Score 38; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 9.e-31;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 KDSADRTTRVDFNAKISIDNFVEINNRVSGAGRKAS 216  
|||||  
DB 111 KDSADRTTRVDFNAKISIDNFVEINNRVSGAGRKAS 148  
|||||

RESULT 123  
O32666  
ID O32666 PRELIMINARY; PRT; 160 AA.  
AC O32666;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP031;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63265; AAB61874.1; -  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17871 MW; 5E9512C7AC382596 CRC64;

Query Match 2.9%; Score 38; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 9.7e-31;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 KDSADRTTRVDFNAKISIDNFVEINNRVSGAGRKAS 216  
|||||  
DB 123 KDSADRTTRVDFNAKISIDNFVEINNRVSGAGRKAS 160  
|||||

RESULT 124  
O32668  
ID O32668 PRELIMINARY; PRT; 160 AA.  
AC O32668;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-HP044;  
RA Beech R.N., Gottke M.O., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63267; AAB61876.1; -  
FT NON\_TER 1  
FT NON\_TER 160  
SQ SEQUENCE 160 AA; 17855 MW; 757D775323884859 CRC64;

Query Match 2.9%; Score 38; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 9.7e-31;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGIRNFTGGDLVNNQKATLRGQNGNSFTSYKDSA 182  
|||||  
DB 89 LSGIRNFTGGDLVNNQKATLRGQNGNSFTSYKDSA 126  
|||||

RESULT 125  
O87486 PRELIMINARY; PRT; 404 AA.  
AC O87486;  
DT 01-NOV-1998 (TRENBLrel. 08, Created)  
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
DE VACUOLATING TOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=539;  
RX MEDLINE=98440809; PubMed-9753641;  
RA Wang H.J., Chang P.C.L., Kuo C.H., Tzeng C.S., Wang W.C.;  
RT "Characterization of the C-terminal domain of Helicobacter pylori  
RT vacuolating toxin and its relationship with extracellular toxin  
RT production.";  
RL Biochem. Biophys. Res. Commun. 250:397-402(1998).  
DR EMBL; AF077940; AAC63522.1; -  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 404 AA; 44023 MW; 0AB6049A935C5159 CRC64;

Query Match 2.9%; Score 38; DB 2; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.2e-30;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 LNNIASLEHKTSGQLTSLSNAMILNRLVNLRRHTN 990  
|||||  
DB 61 LNNIASLEHKTSGQLTSLSNAMILNRLVNLRRHTN 98  
|||||

RESULT 126  
O87485 PRELIMINARY; PRT; 406 AA.  
AC O87485;  
DT 01-NOV-1998 (TRENBLrel. 08, Created)  
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
DE VACUOLATING TOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=166;  
RX MEDLINE=98440809; PubMed-9753641;  
RA Wang H.J., Chang P.C.L., Kuo C.H., Tzeng C.S., Wang W.C.;

RT "Characterization of the C-terminal domain of Helicobacter pylori  
RT vacuolating toxin and its relationship with extracellular toxin  
RT production.";  
RL Biochem. Biophys. Res. Commun. 250:397-402(1998).  
DR EMBL; AF077939; AAC63521.1; -  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 406 AA; 44294 MW; B195240CF757266C CRC64;

Query Match 2.9%; Score 38; DB 2; Length 406;  
Best Local Similarity 100.0%; Pred. No. 2.2e-30;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 LNNIASLEHKTSGQLTSLSNAMILNRLVNLRRHTN 990  
|||||  
DB 63 LNNIASLEHKTSGQLTSLSNAMILNRLVNLRRHTN 100  
|||||

RESULT 127  
Q9ZHT8 PRELIMINARY; PRT; 1291 AA.  
AC Q9ZHT8;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F64;  
RX MEDLINE=98453456; PubMed-9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and  
RT noncytotoxic Helicobacter pylori.";  
RL J. Infect. Dis. 178:1391-1398(1998).  
DR EMBL; AF049647; AAD04282.1; -  
SQ SEQUENCE 1291 AA; 139395 MW; 4B5785B5BEE7934 CRC64;

Query Match 2.9%; Score 38; DB 2; Length 1291;  
Best Local Similarity 100.0%; Pred. No. 6.2e-30;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 LNNIASLEHKTSGQLTSLSNAMILNRLVNLRRHTN 990  
|||||  
DB 948 LNNIASLEHKTSGQLTSLSNAMILNRLVNLRRHTN 985  
|||||

RESULT 128  
Q9R961 PRELIMINARY; PRT; 1291 AA.  
AC Q9R961;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F65;  
RX MEDLINE=98453456; PubMed-9780260;  
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
RA Kuriyama M.;  
RT "Full-length sequence analysis of the vacA gene from cytotoxic and

```

RT noncytotoxic Helicobacter pylori."
RL J. Infect. Dis. 178:1391-1398(1998).
DR EMBL: AF049636; RAD04272.1; -
SQ SEQUENCE 1291 AA; 139617 MW; A6A93343EE7C6448 CRC64;

Query Match          2.9%; Score 38; DB 2; Length 1291;
Best Local Similarity 100.0%; Pred. No. 6.2e-30;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 LNNIASLEHKTSGLOTLISLNASMILNSRLVNLSSRRHTN 990
Db 948 LNNIASLEHKTSGLOTLISLNASMILNSRLVNLSSRRHTN 985

RESULT 129
Q9ZHT4          PRELIMINARY; PRT; 1298 AA.
AC Q9ZHT4;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE VACUOLATING CYTOXIN PRECURSOR.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=F70;
RX MEDLINE=98453456; PubMed=9780260;
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
RA Kuriyama M.;
RT "Full-length sequence analysis of the vacA gene from cytotoxic and
RT noncytotoxic Helicobacter pylori.";
RL J. Infect. Dis. 178:1391-1398(1998).
DR EMBL: AF049650; RAD04286.1; -
SQ SEQUENCE 1298 AA; 140250 MW; A2F3B9003256F7DE CRC64;

Query Match          2.9%; Score 38; DB 2; Length 1298;
Best Local Similarity 100.0%; Pred. No. 6.2e-30;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 LNNIASLEHKTSGLOTLISLNASMILNSRLVNLSSRRHTN 990
Db 955 LNNIASLEHKTSGLOTLISLNASMILNSRLVNLSSRRHTN 992

RESULT 130
Q9LBB4          PRELIMINARY; PRT; 131 AA.
AC Q9LBB4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE VACUOLATING CYTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CHN3132B;
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,
RA Rayat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF050335; AAF26517.1; -
FT NON_TER 1
FT NON_TER 131

```

```

SQ SEQUENCE 131 AA; 14587 MW; B8C9A6D5D335F4B0 CRC64;

Query Match          2.9%; Score 37; DB 2; Length 131;
Best Local Similarity 100.0%; Pred. No. 8.9e-30;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 761 EOPKERLALYNNNRMDTCVVRNTDDIKACGMAIGDQ 797
Db 93 EOPKERLALYNNNRMDTCVVRNTDDIKACGMAIGDQ 129

RESULT 131
Q9R846          PRELIMINARY; PRT; 148 AA.
AC Q9R846;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE VACUOLATING CYTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CC48;
RX MEDLINE=98445420; PubMed=9770535;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL: AJ009433; CAA08708.1; -
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16661 MW; 5EEDBA1820425F29 CRC64;

Query Match          2.9%; Score 37; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 9.9e-30;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 TGGDLDVNNQKATRLGQFNGNSFTSYKDSADTRTV 188
Db 84 TGGDLDVNNQKATRLGQFNGNSFTSYKDSADTRTV 120

RESULT 132
Q32656          PRELIMINARY; PRT; 160 AA.
AC Q32656;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE VACUOLATING CYTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HP001;
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U63255; AAB61864.1; -
FT NON_TER 1
FT NON_TER 160
SQ SEQUENCE 160 AA; 17888 MW; EFCDD20C5C7AD7CDF CRC64;

Query Match          2.9%; Score 37; DB 2; Length 160;

```

Best Local Similarity 100.0%; Pred. No. 1.1e-29;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 TGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
Db 96 TGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 132

## RESULT 133:

O32678  
ID O32678 PRELIMINARY; PRT; 160 AA.  
AC O32678;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP079;  
RA Beech R.N., Gottke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63277; AAB61886.1; -  
FT NON\_TER 1 160  
FT NON\_TER 160 160  
SQ SEQUENCE 160 AA; 17888 MW; 657751D40DD5DFC3 CRC64;

Query Match 2.9%; Score 37; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 1.1e-29;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 TGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 188  
Db 96 TGGDLVNMOKATRLGQFNGNSFTSYKDSADRTTRV 132

## RESULT 134

O53040  
ID O53040 PRELIMINARY; PRT; 206 AA.  
AC O53040;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=049;  
RA Strobel S., Bereswill S., Ballig P., Allgaier P., Sonntag H.G.,  
RA Kist M.;  
RL J. Clin. Microbiol. 0:0-0(0).  
DR EMBL; Y14744; CAA75035.1; -  
FT NON\_TER 1 206  
FT NON\_TER 206 206  
SQ SEQUENCE 206 AA; 22775 MW; 9953AFFE478D8057 CRC64;

Query Match 2.9%; Score 37; DB 2; Length 206;  
Best Local Similarity 100.0%; Pred. No. 1.3e-29;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 640 QSPWGTSLKMFNNLTGQNAVMDYSQFSNLTIQGDF 676  
Db 74 QSPWGTSLKMFNNLTGQNAVMDYSQFSNLTIQGDF 110

## RESULT 135

Q9S312  
ID Q9S312 PRELIMINARY; PRT; 148 AA.  
AC Q9S312;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BO265;  
RX MEDLINE=99255683; PubMed=10320570;  
RA Achtmann M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,  
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;  
RT "Recombination and clonal groupings within Helicobacter pylori from  
different geographic regions."  
RL Mol. Microbiol. 32:459-470(1999).  
DR EMBL; AJ239559; CAB37371.1; -  
FT NON\_TER 1 148  
FT NON\_TER 148 148  
SQ SEQUENCE 148 AA; 16502 MW; 135DA92494D08BE0 CRC64;

Query Match 2.8%; Score 36; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.1e-28;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSLGRNFTGGDLVNMOKATRLGQFNGNSFTSYKD 180  
Db 77 LSLGRNFTGGDLVNMOKATRLGQFNGNSFTSYKD 112

## RESULT 136

Q9R2Y5  
ID Q9R2Y5 PRELIMINARY; PRT; 35 AA.  
AC Q9R2Y5;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACA PROTEIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=US123, NL5825, NL4819, PO30, NL4705, NL4330, CH5, AND US178;  
RA van Doorn L.J.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=US123, NL5825, NL4819, PO30, NL4705, NL4330, CH5, AND US178;  
RX MEDLINE=98371099; PubMed=9705399;  
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,  
RA Atherton J.C., Blaser M.J., Quint W.G.;  
RT "Expanding allelic diversity of Helicobacter pylori vacA."  
RL J. Clin. Microbiol. 36:2597-2603(1998).  
DR EMBL; AJ390705; CAB64550.1; -  
DR EMBL; AJ390699; CAB64418.1; -  
DR EMBL; AJ390703; CAB64420.1; -  
DR EMBL; AJ390706; CAB64422.1; -  
DR EMBL; AJ390707; CAB64423.1; -  
DR EMBL; AJ390708; CAB64424.1; -  
DR EMBL; AJ390709; CAB64425.1; -  
DR EMBL; AJ390700; CAB64549.1; -

FT NON\_TER 1 1  
 FT NON\_TER 35 35  
 SQ SEQUENCE 35 AA; 3489 MW; CB697D8BFF819054 CRC64;

Query Match 2.7%; Score 35; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-28;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 PLVSLALVGLVSVITPOOSHAAFFTTVIIIAIVGG 47  
 Db 1 PLVSLALVGLVSVITPOOSHAAFFTTVIIIAIVGG 35

RESULT 137  
 Q9REB2 ID Q9REB2 PRELIMINARY; PRT; 96 AA.  
 AC Q9REB2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE VACA PROTEIN (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NL4602;  
 RA van Doorn L.J.;  
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NL4602;  
 RX MEDLINE=98371099; PubMed=9705339;  
 RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,  
 RA Atherton J.C., Blaser M.J., Quint W.G.;  
 RT "Expanding allelic diversity of Helicobacter pylori vacA.";  
 RL J. Clin. Microbiol. 36:2597-2603(1998).  
 DR EMBL; AJ390597; CAB64367.1;  
 FT NON\_TER 1 1  
 FT NON\_TER 96 96  
 SQ SEQUENCE 96 AA; 10177 MW; 2F17E39FA79208B9 CRC64;

Query Match 2.7%; Score 35; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-28;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VKTNGISVGVEYTHFSDIGSQSRINTVRLTGTGRS 594  
 Db 54 VKTNGISVGVEYTHFSDIGSQSRINTVRLTGTGRS 88

RESULT 138  
 Q9LB84 ID Q9LB84 PRELIMINARY; PRT; 125 AA.  
 AC Q9LB84;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CHN4590A;  
 RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
 RA Rayrat J.M., Xu G., Rappuoli R.;

RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF050366; AAF26548.1;  
 FT NON\_TER 1 1  
 FT NON\_TER 125 125  
 SQ SEQUENCE 125 AA; 13259 MW; F501C5C3C6DFAEBB CRC64;

Query Match 2.7%; Score 35; DB 2; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 1e-27;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 SOSGAKNDKKNESAKNDKQESSQNSNTQVINPPNS 374  
 Db 70 SOSGAKNDKKNESAKNDKQESSQNSNTQVINPPNS 104

RESULT 139  
 Q9LBC9 ID Q9LBC9 PRELIMINARY; PRT; 127 AA.  
 AC Q9LBC9;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPM326;  
 RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
 RA Rayrat J.M., Xu G., Rappuoli R.;  
 RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF050318; AAF26500.1;  
 FT NON\_TER 1 1  
 FT NON\_TER 127 127  
 SQ SEQUENCE 127 AA; 13405 MW; 93D027F8459F2B05 CRC64;

Query Match 2.7%; Score 35; DB 2; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 1e-27;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 KNDKNESAKNDKQESSQNSNTQVINPPNSAQKTE 379  
 Db 75 KNDKNESAKNDKQESSQNSNTQVINPPNSAQKTE 109

RESULT 140  
 Q9LBB3 ID Q9LBB3 PRELIMINARY; PRT; 144 AA.  
 AC Q9LBB3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CHN3285D;  
 RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
 RA Rayrat J.M., Xu G., Rappuoli R.;  
 RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF050336; AAF26518.1;



```

FT NON_TER 1
FT NON_TER 144
SQ SEQUENCE 144 AA; 15339 MW; 4C43759C05D8D8A9 CRC64;

Query Match
Best Local Similarity 100.0%; DB 2; Length 144;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 SQSGAKNDKNSAKNDKQSSQNNSTQVNPNS 374
    |||||||
DB 70 SQSGAKNDKNSAKNDKQSSQNNSTQVNPNS 104

RESULT 141
Q9L7T1 PRELIMINARY; PRT; 216 AA.
AC Q9L7T1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA48;
RX MEDLINE-20270153; PubMed-10809703;
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,
RA Nair G.B., Berg D.E.;
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,
RT India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL; AF220110; AAF42854.1;
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 23849 MW; 6052CABCE0429A0 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; DB 2; Length 216;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 560 VKTNGISVGEYTHFSEDIGSQSRINTVRLTGTTRS 594
    |||||||
DB 8 VKTNGISVGEYTHFSEDIGSQSRINTVRLTGTTRS 42

RESULT 142
Q9L7S9 PRELIMINARY; PRT; 216 AA.
AC Q9L7S9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA48;
RX MEDLINE-20270153; PubMed-10809703;
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,
RA Nair G.B., Berg D.E.;
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,
RT India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL; AF220110; AAF42854.1;
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 23849 MW; 6052CABCE0429A0 CRC64;

```

```

RT India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL; AF220112; AAF42856.1;
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 23911 MW; 6D978460B70482B2 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; DB 2; Length 216;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 560 VKTNGISVGEYTHFSEDIGSQSRINTVRLTGTTRS 594
    |||||||
DB 8 VKTNGISVGEYTHFSEDIGSQSRINTVRLTGTTRS 42

```

```

RESULT 143
Q9L7S7 PRELIMINARY; PRT; 216 AA.
AC Q9L7S7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA89;
RX MEDLINE-20270153; PubMed-10809703;
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,
RA Nair G.B., Berg D.E.;
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,
RT India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL; AF220114; AAF42858.1;
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 23851 MW; 6FEDE175DAC487A0 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; DB 2; Length 216;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 560 VKTNGISVGEYTHFSEDIGSQSRINTVRLTGTTRS 594
    |||||||
DB 8 VKTNGISVGEYTHFSEDIGSQSRINTVRLTGTTRS 42

```

```

RESULT 144
Q9L7S5 PRELIMINARY; PRT; 216 AA.
AC Q9L7S5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA227;
RX MEDLINE-20270153; PubMed-10809703;
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,

```

```

RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,
RA Nair G.B., Berg D.E.;
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,
RT India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL: AF220116; AAF42860.1; -
FT NON_TER 216 216
SQ SEQUENCE 216 AA; 23835 MW; 9B43202F93FCCE79 CRC64;

Query Match 2.7%; Score 35; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.7e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VKTNGISVGVEYTHFSEDIGSQSRINTVRLTGTGRS 594
Db 8 VKTNGISVGVEYTHFSEDIGSQSRINTVRLTGTGRS 42

RESULT 145
Q9L7S4
ID Q9L7S4 PRELIMINARY; PRT; 216 AA.
AC Q9L7S4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA230;
RX MEDLINE-20270153; PubMed-10809703;
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,
RA Nair G.B., Berg D.E.;
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,
RT India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL: AF220117; AAF42861.1; -
FT NON_TER 216 216
SQ SEQUENCE 216 AA; 23848 MW; D509EB2493962C79 CRC64;

Query Match 2.7%; Score 35; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.7e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VKTNGISVGVEYTHFSEDIGSQSRINTVRLTGTGRS 594
Db 8 VKTNGISVGVEYTHFSEDIGSQSRINTVRLTGTGRS 42

RESULT 146
Q9L7S1
ID Q9L7S1 PRELIMINARY; PRT; 223 AA.
AC Q9L7S1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-INDIA100;
RX MEDLINE-20270153; PubMed-10809703;
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,
RA Nair G.B., Berg D.E.;
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,
RT India.";
RL J. Bacteriol. 182:3219-3227(2000).
DR EMBL: AF220120; AAF42864.1; -
FT NON_TER 223 223
SQ SEQUENCE 223 AA; 24812 MW; C0A8D0727D0A1B3F CRC64;

Query Match 2.7%; Score 35; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.7e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VKTNGISVGVEYTHFSEDIGSQSRINTVRLTGTGRS 594
Db 12 VKTNGISVGVEYTHFSEDIGSQSRINTVRLTGTGRS 46

RESULT 147
Q9RE81
ID Q9RE81 PRELIMINARY; PRT; 96 AA.
AC Q9RE81;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-US178;
RA van Doorn L.J.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-US178;
RX MEDLINE-98371099; PubMed-9705399;
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,
RA Atherton J.C., Blaser M.J., Quint W.G.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL: AJ390598; CAB64274.1; -
FT NON_TER 96 96
SQ SEQUENCE 96 AA; 10260 MW; 70AAE4ABDD51C427 CRC64;

Query Match 2.6%; Score 34; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 8.9e-27;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 561 KTGNGISVGVEYTHFSEDIGSQSRINTVRLTGTGRS 594
Db 55 KTGNGISVGVEYTHFSEDIGSQSRINTVRLTGTGRS 88

RESULT 148
Q9LBC5
ID Q9LBC5 PRELIMINARY; PRT; 128 AA.
AC Q9LBC5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.

```

OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=G21;  
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,  
RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050322; AAF26504.1; -  
FT NON\_TER 1  
FT NON\_TER 128 128  
SQ SEQUENCE 128 AA; 13550 MW; 704026E3849F5F5C CRC64;

Query Match 2.5%; Score 34; DB 2; Length 128;  
Best Local Similarity 100.0%; Pred. No. 1.2e-26;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQSLNIAPPEGGYKDKPNTPSQSG 343  
Db 40 IGTLDLWQSLNIAPPEGGYKDKPNTPSQSG 73

RESULT 149  
QYLB69 PRELIMINARY; PRT; 112 AA.  
AC QYLB69  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CHN5038C;  
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,  
RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050383; AAF26565.1; -  
FT NON\_TER 1  
FT NON\_TER 112 112  
SQ SEQUENCE 112 AA; 11800 MW; 3A7DFC28F29BA2E1 CRC64;

Query Match 2.5%; Score 33; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.1e-25;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 KKTNICTLDLWQSLNIAPPEGGYKDKPNT 338  
Db 20 KKTNICTLDLWQSLNIAPPEGGYKDKPNT 52

RESULT 150  
QYLB66 PRELIMINARY; PRT; 127 AA.  
AC QYLB66  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.

OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CHN1811H;  
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,  
RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050333; AAF26515.1; -  
FT NON\_TER 1  
FT NON\_TER 127 127  
SQ SEQUENCE 127 AA; 14127 MW; 271AC3E8286B2802 CRC64;

Query Match 2.5%; Score 33; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 1.3e-25;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 ERLALYNNNNRMDTCVVRNTDDIKACGMAIGDQ 797  
Db 93 ERLALYNNNNRMDTCVVRNTDDIKACGMAIGDQ 125

RESULT 151  
QYLB79 PRELIMINARY; PRT; 140 AA.  
AC QYLB79  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CHN4611A;  
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,  
RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050372; AAF26554.1; -  
FT NON\_TER 1  
FT NON\_TER 140 140  
SQ SEQUENCE 140 AA; 15390 MW; 94742A1527C2BBDB CRC64;

Query Match 2.5%; Score 33; DB 2; Length 140;  
Best Local Similarity 100.0%; Pred. No. 1.4e-25;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 ERLALYNNNNRMDTCVVRNTDDIKACGMAIGDQ 797  
Db 106 ERLALYNNNNRMDTCVVRNTDDIKACGMAIGDQ 138

RESULT 152  
QYLB68 PRELIMINARY; PRT; 144 AA.  
AC QYLB68  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=CHN5038C;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RT "Allergic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050384; AAF26566.1; -
FT NON_TER 1
FT NON_TER 144
SQ SEQUENCE 144 AA; 15723 MW; E02FF583DED87B70 CRC64;

Query Match 2.5%; Score 33; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.4e-25;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 ERLALYNNNRMDTCVVRNTDDIKACGMAIGDQ 797
DB 110 ERLALYNNNRMDTCVVRNTDDIKACGMAIGDQ 142

RESULT 153
Q9LB97
ID Q9LB97 PRELIMINARY; PRT; 145 AA.
AC Q9LB97;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN3554A;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RT "Allergic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050353; AAF26535.1; -
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 15853 MW; BFA3A5BE72188CD4 CRC64;

Query Match 2.5%; Score 33; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.4e-25;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 ERLALYNNNRMDTCVVRNTDDIKACGMAIGDQ 797
DB 111 ERLALYNNNRMDTCVVRNTDDIKACGMAIGDQ 143

RESULT 154
Q9S3H9
ID Q9S3H9 PRELIMINARY; PRT; 148 AA.
AC Q9S3H9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=88-39;
RX MEDLINE=99255683; PubMed-10320570;
RA Achtmann M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,

```

```

RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239566; CAB37378.1; -
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16666 MW; E57CAD38C6DBESF4 CRC64;

Query Match 2.5%; Score 33; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.4e-25;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFTGGDLVDVNMOKATRLRGQFNGNSFTS 177
DB 77 LSGLRNFTGGDLVDVNMOKATRLRGQFNGNSFTS 109

RESULT 155
Q9S3H7
ID Q9S3H7 PRELIMINARY; PRT; 148 AA.
AC Q9S3H7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-42;
RX MEDLINE=99255683; PubMed-10320570;
RA Achtmann M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239568; CAB37380.1; -
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16644 MW; 0C0C33459FDC07E0 CRC64;

Query Match 2.5%; Score 33; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.4e-25;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSGLRNFTGGDLVDVNMOKATRLRGQFNGNSFTS 177
DB 77 LSGLRNFTGGDLVDVNMOKATRLRGQFNGNSFTS 109

RESULT 156
Q9LBA1
ID Q9LBA1 PRELIMINARY; PRT; 148 AA.
AC Q9LBA1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN3503C;
RX MEDLINE=99255683; PubMed-10320570;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,

```

RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050349; AAF26531.1;  
FT NON\_TER 148 148  
SQ SEQUENCE 148 AA; 16174 MW; 6D5D5B214673F547 CRC64;

Query Match 2.5%; Score 33; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.4e-25;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 ERLALNNNNRMDTCVVRNTDDIKACGMAIGDQ 797  
Db 114 ERLALNNNNRMDTCVVRNTDDIKACGMAIGDQ 146

RESULT 157  
O32680  
ID O32680 PRELIMINARY; PRT; 160 AA.

AC O32680;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HP110;  
RA Beech R.N., Gotke M.U., Fallone C.A., Loo V., Barkun A.N.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63279; AAB61888.1;  
FT NON\_TER 1 1  
SQ SEQUENCE 160 AA; 17739 MW; 468FBAAE62EBE79B CRC64;

Query Match 2.5%; Score 33; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 1.5e-25;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 RTRVDFNAKNISIDNFEVNNRVGSGAGRKAS 216  
Db 128 RTRVDFNAKNISIDNFEVNNRVGSGAGRKAS 160

RESULT 158  
Q9R356  
ID Q9R356 PRELIMINARY; PRT; 210 AA.

AC Q9R356;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=T-47, AND T-12;  
RA Lin C.W.;  
RT "Helicobacter pylori Taiwanese isolates.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF091833; AAD23370.1;  
DR EMBL; AF091824; AAD23361.1;  
DR InterPro; IPR000897; -.

DR ProDom; PD000819; -; 1.  
FT NON\_TER 210 210  
SQ SEQUENCE 210 AA; 23075 MW; A7723A29B93FED92 CRC64;

Query Match 2.5%; Score 33; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 2e-25;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSLGRNFTGGDLVNNQKATRLRGQFNGNSFTS 177  
Db 145 LSLGRNFTGGDLVNNQKATRLRGQFNGNSFTS 177

RESULT 159  
Q9F9G2  
ID Q9F9G2 PRELIMINARY; PRT; 861 AA.

AC Q9F9G2;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=85;  
RA Kuo C.H., Wang W.C.;  
RT "Genetic and functional comparison of Helicobacter pylori vacuolating toxin between sla/mlr and sla/m2 isolates.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF195013; AAG28441.1;  
FT NON\_TER 1 1  
SQ SEQUENCE 861 AA; 92773 MW; D47E25A5274827D0 CRC64;

Query Match 2.5%; Score 32; DB 2; Length 861;  
Best Local Similarity 100.0%; Pred. No. 7.6e-24;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGNVMGRLQYVGAYLAPSYSTINTSKV 281  
Db 217 SVKLMGNVMGRLQYVGAYLAPSYSTINTSKV 248

RESULT 160  
Q9L810  
ID Q9L810 PRELIMINARY; PRT; 78 AA.

AC Q9L810;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=INDIA21;  
RX MEDLINE-20270153; PubMed-10809703;  
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,  
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,  
RA Nair G.B., Berg D.E.;  
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta, India.";  
RL J. Bacteriol. 182:3219-3227(2000).  
DR EMBL; AF217732; AAF42843.1; -.

```
FT NON_TER 1 1
FT NON_TER 78 78
SQ SEQUENCE 78 AA; 8093 MW; B97A3AA21C382363 CRC64;

Query Match 2.4%; Score 31; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 9.8e-24;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 SHAAFTTIIIPAVGGIATGTAAGTVSGLL 61
DB 23 SHAAFTTIIIPAVGGIATGTAAGTVSGLL 53

RESULT 161
Q9KI22 PRELIMINARY; PRT; 118 AA.
ID Q9KI22
AC Q9KI22;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE M VACUOLATING CYTOTOXIN A (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RA Ji X., Telford J.L.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF244143; AAF78033.1;
FT NON_TER 1 1
FT NON_TER 118 118
SQ SEQUENCE 118 AA; 12853 MW; 3CE8B1557DF52DA6 CRC64;

Query Match 2.4%; Score 31; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 ERLALYNNNRMDTCVVRTDDIKACGMAIG 795
DB 88 ERLALYNNNRMDTCVVRTDDIKACGMAIG 118

RESULT 162
Q9R836 PRELIMINARY; PRT; 148 AA.
ID Q9R836
AC Q9R836;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DE VACUOLATING CYTOTOXIN A (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=FE5;
RC MEDLINE=98445420; PubMed=9770535;
RX Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009446; CAA08721.1;
FT NON_TER 1 1
FT NON_TER 148 148
SQ SEQUENCE 148 AA; 16632 MW; 24FE9B3B974E1B86 CRC64;
```

```
Query Match 2.4%; Score 31; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 NFTGGDLVNMOKATRLRGQFNGNSFTSYKD 180
DB 82 NFTGGDLVNMOKATRLRGQFNGNSFTSYKD 112

RESULT 163
Q9LAJ8 PRELIMINARY; PRT; 214 AA.
ID Q9LAJ8
AC Q9LAJ8;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE VACUOLATING CYTOTOXIN A (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RA Pan S.K., Kuo C.H., Su I.J., Su R., Wang W.C.;
RT "Characterization of Helicobacter pylori strains isolated from
RT Helicobacter pylori-infected spouses in Taiwan.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF132301; AAF61324.1;
FT NON_TER 1 1
FT NON_TER 214 214
SQ SEQUENCE 214 AA; 23810 MW; E9CD4751C1A75716 CRC64;

Query Match 2.4%; Score 31; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-23;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNYFDARNIKNVEITNKLAFGPGSPWGTS 647
DB 71 PNYFDARNIKNVEITNKLAFGPGSPWGTS 101

RESULT 164
Q9R9B0 PRELIMINARY; PRT; 244 AA.
ID Q9R9B0
AC Q9R9B0;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DE VACUOLATING CYTOTOXIN A (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang H.J., Kuo C.H., Yeh A.M., Chang C.L., Wang W.C.;
RC STRAIN=61;
RT "Vacuolating toxin production in clinical isolates of Helicobacter
RT pylori with different vacA genotypes.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030401; AAD01911.1;
FT NON_TER 1 1
FT NON_TER 244 244
SQ SEQUENCE 244 AA; 26596 MW; 18FASD241D61DA0B CRC64;

Query Match 2.4%; Score 31; DB 2; Length 244;
```

```

RESULT 166
Q9RE80
ID Q9RE80 PRELIMINARY; PRT; 96 AA.
AC Q9RE80;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_taxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PF9019;
RA van Doorn L.J.;
RRL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PF9019;
RX MEDLINE=98371099; PubMed=9705399;
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,
RA Atherton J.C., Blaser M.J., Quint W.G.;

```

```
RC STRAIN-JA32;
RA van Doorn L.J.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-JA32;
RX MEDLINE=98371099; PubMed=97053399;
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,
RT Atherton J.C., Blaser M.J., Quint W.G.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390608; CAB64283.1; -
DR EMBL; AJ390607; CAB64282.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10267 MW; CA74D55DE09F138E CRC64;

Query Match 2.3%; Score 30; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 535 KVNINKLITASTNVAVKFNINELIVKTN 564
Db 29 KVNINKLITASTNVAVKFNINELIVKTN 58

RESULT 169
QYR2R3 ID Q9R2R3 PRELIMINARY; PRT; 96 AA.
AC Q9R2R3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PE9012-1, AND PO12;
RA van Doorn L.J.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PE9012-1, AND PO12;
RX MEDLINE=98371099; PubMed=97053399;
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,
RT Atherton J.C., Blaser M.J., Quint W.G.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390601; CAB64277.1; -
DR EMBL; AJ390599; CAB64275.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10210 MW; EC19D55DECA32382 CRC64;

Query Match 2.3%; Score 30; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 535 KVNINKLITASTNVAVKFNINELIVKTN 564
Db 29 KVNINKLITASTNVAVKFNINELIVKTN 58

RESULT 170
QYLB49 ID Q9LB49 PRELIMINARY; PRT; 130 AA.
AC Q9LB49;
```

```
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-G104;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RT Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050403; AAF26585.1; -
FT NON_TER 1
FT NON_TER 130
SQ SEQUENCE 130 AA; 14343 MW; 9BE567D1F0540AA5 CRC64;

Query Match 2.3%; Score 30; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.7e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 529 FSGVTDKVNINKLITASTNVAVKFNINEL 558
Db 27 FSGVTDKVNINKLITASTNVAVKFNINEL 56

RESULT 171
QYK2M6 ID Q9K2M6 PRELIMINARY; PRT; 214 AA.
AC Q9K2M6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE VACUOLATING CYTOTOXIN A (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-V296, AND V294;
RA Pan S.K., Kuo C.H., Su I.J., Su R., Wang W.C.;
RT "Characterization of Helicobacter pylori strains isolated from
Helicobacter pylori-infected spouses in Taiwan."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF132306; AAF61329.1; -
DR EMBL; AF132305; AAF61328.1; -
DR InterPro; IPR001809; -
DR ProDom; PD001127; -; 1.
FT NON_TER 1
FT NON_TER 214
SQ SEQUENCE 214 AA; 23404 MW; D323AF0F5D453EB3 CRC64;

Query Match 2.3%; Score 30; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.6e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PWNFYDARNIKNVETNKLAFPGQSPWGT 646
Db 71 PWNFYDARNIKNVETNKLAFPGQSPWGT 100

RESULT 172
QYLB70 ID Q9LB70 PRELIMINARY; PRT; 216 AA.
AC Q9LB70;
```



DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=INDIA19;  
RX MEDLINE=20270153; PubMed=10809703;  
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,  
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,  
RA Nair G.B., Berg D.E.;  
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,  
RT India.";  
RL J. Bacteriol. 182:3219-3227(2000).  
DR EMBL; AF220111; AAF42855.1; -;  
FT NON\_TER 1 216  
SQ SEQUENCE 216 AA; 23779 MW; 129EA0D9004252BA CRC64;  
  
Query Match 2.3%; Score 30; DB 2; Length 216;  
Best Local Similarity 100.0%; Pred. No. 2.7e-22;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 617 PNIFYDARNIKNVEITNKLAFGPQGS PWGT 646  
Db 65 PNIFYDARNIKNVEITNKLAFGPQGS PWGT 94  
  
RESULT 173  
Q9L7S8 PRELIMINARY; PRT; 216 AA.  
AC Q9L7S8;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=INDIA66;  
RX MEDLINE=20270153; PubMed=10809703;  
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,  
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,  
RA Nair G.B., Berg D.E.;  
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,  
RT India.";  
RL J. Bacteriol. 182:3219-3227(2000).  
DR EMBL; AF220113; AAF42857.1; -;  
FT NON\_TER 1 216  
SQ SEQUENCE 216 AA; 23793 MW; CDB495F5B7617BA2 CRC64;

Query Match 2.3%; Score 30; DB 2; Length 216;  
Best Local Similarity 100.0%; Pred. No. 2.7e-22;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 617 PNIFYDARNIKNVEITNKLAFGPQGS PWGT 646  
Db 65 PNIFYDARNIKNVEITNKLAFGPQGS PWGT 94  
  
RESULT 174

Q9L7S6 PRELIMINARY; PRT; 216 AA.  
AC Q9L7S6;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=INDIA226;  
RX MEDLINE=20270153; PubMed=10809703;  
RA Mukhopadhyay A.K., Kersulyte D., Jeong J.Y., Datta S., Ito Y.,  
RA Chowdhury A., Chowdhury S., Santra A., Bhattacharya S.K., Azuma T.,  
RA Nair G.B., Berg D.E.;  
RT "Distinctiveness of genotypes of Helicobacter pylori in Calcutta,  
RT India.";  
RL J. Bacteriol. 182:3219-3227(2000).  
DR EMBL; AF220115; AAF42859.1; -;  
FT NON\_TER 1 216  
SQ SEQUENCE 216 AA; 23843 MW; CFE07BC3F84E74A3 CRC64;  
  
Query Match 2.3%; Score 30; DB 2; Length 216;  
Best Local Similarity 100.0%; Pred. No. 2.7e-22;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 617 PNIFYDARNIKNVEITNKLAFGPQGS PWGT 646  
Db 65 PNIFYDARNIKNVEITNKLAFGPQGS PWGT 94  
  
RESULT 175  
Q9X412 PRELIMINARY; PRT; 240 AA.  
AC Q9X412;  
DT 01-NOV-1999 (TReMBLrel. 12, Created)  
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=T-47;  
RA Lin C.W.;  
RT "Helicobacter pylori Taiwanese isolates.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF091819; AAD23356.1; -;  
FT NON\_TER 1 240  
SQ SEQUENCE 240 AA; 26644 MW; 5097F98B5F3037AA CRC64;  
  
Query Match 2.3%; Score 30; DB 2; Length 240;  
Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 617 PNIFYDARNIKNVEITNKLAFGPQGS PWGT 646  
Db 77 PNIFYDARNIKNVEITNKLAFGPQGS PWGT 106  
  
RESULT 176  
O34107

```

ID O34107 PRELIMINARY; PRT; 244 AA.
AC O34107;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE STRAIN F84 VACA (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F84;
RX MEDLINE=97339580; PubMed=9196179;
RA Ito Y., Azuma T., Ito S., Miyaji H., Hirai M., Yamazaki Y., Sato F.,
RA Kato T., Kohli Y., Kuriyama M.;
RT "Analysis and typing of the vca gene from caga-positive strains of
RT Helicobacter pylori isolated in Japan.";
RL J. Clin. Microbiol. 35:1710-1714(1997).
DR EMBL: U91575; AAB66698.1; -
FT NON_TER 1 1
FT NON_TER 244 244
SQ SEQUENCE 244 AA; 26572 MW; 02D649A40BB55AA6 CRC64;

Query Match 2.3%; Score 30; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 3e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNNYFDARNIKNVEITNKLAFGPGQSPWGT 646
|||||
DB 101 PNNYFDARNIKNVEITNKLAFGPGQSPWGT 130

RESULT 177
O34108 PRELIMINARY; PRT; 244 AA.
AC O34108;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOXIN.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F71;
RX MEDLINE=97339580; PubMed=9196179;
RA Ito Y., Azuma T., Ito S., Miyaji H., Hirai M., Yamazaki Y., Sato F.,
RA Kato T., Kohli Y., Kuriyama M.;
RT "Analysis and typing of the vca gene from caga-positive strains of
RT Helicobacter pylori isolated in Japan.";
RL J. Clin. Microbiol. 35:1710-1714(1997).
DR EMBL: AF071096; AAC77451.1; -
SQ SEQUENCE 244 AA; 26613 MW; 1353433A6F5CA271 CRC64;

Query Match 2.3%; Score 30; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 3e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNNYFDARNIKNVEITNKLAFGPGQSPWGT 646
|||||
DB 101 PNNYFDARNIKNVEITNKLAFGPGQSPWGT 130

RESULT 178
O34109 PRELIMINARY; PRT; 244 AA.
AC O34109;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOXIN.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F71;
RX MEDLINE=97339580; PubMed=9196179;
RA Ito Y., Azuma T., Ito S., Miyaji H., Hirai M., Yamazaki Y., Sato F.,
RA Kato T., Kohli Y., Kuriyama M.;
RT "Analysis and typing of the vca gene from caga-positive strains of
RT Helicobacter pylori isolated in Japan.";
RL J. Clin. Microbiol. 35:1710-1714(1997).
DR EMBL: AF071096; AAC77451.1; -
SQ SEQUENCE 244 AA; 26613 MW; 1353433A6F5CA271 CRC64;

Query Match 2.3%; Score 30; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 3e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNNYFDARNIKNVEITNKLAFGPGQSPWGT 646
|||||
DB 101 PNNYFDARNIKNVEITNKLAFGPGQSPWGT 130

RESULT 179
O34110 PRELIMINARY; PRT; 83 AA.
AC O34110;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN4099A;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vca gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF050362; AAF26544.1; -
FT NON_TER 1 1
FT NON_TER 83 83
SQ SEQUENCE 83 AA; 8716 MW; 3CAED4C908AA0502 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 ICTLDLWQSAGLNIAPPGGYKDKPNNT 338
|||||
DB 40 ICTLDLWQSAGLNIAPPGGYKDKPNNT 68

RESULT 180
O34111 PRELIMINARY; PRT; 89 AA.
AC O34111;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

```

```

AC O92IE5;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACUOLATING CYTOXIN A (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=184;
RX Wang H.J., Kuo C.H., Yeh A.M., Chang C.L., Wang W.C.;
RT "Vaccinating toxin production in clinical isolates of Helicobacter
RT pylori with different vca genotypes.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF030402; AAD01912.1; -
FT NON_TER 1 1
FT NON_TER 244 244
SQ SEQUENCE 244 AA; 26736 MW; 0F50C0E98F4C2090 CRC64;

Query Match 2.3%; Score 30; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 3e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PNNYFDARNIKNVEITNKLAFGPGQSPWGT 646
|||||
DB 101 PNNYFDARNIKNVEITNKLAFGPGQSPWGT 130

RESULT 179
O9LB88 PRELIMINARY; PRT; 83 AA.
AC O9LB88;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN4099A;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vca gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF050362; AAF26544.1; -
FT NON_TER 1 1
FT NON_TER 83 83
SQ SEQUENCE 83 AA; 8716 MW; 3CAED4C908AA0502 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 ICTLDLWQSAGLNIAPPGGYKDKPNNT 338
|||||
DB 40 ICTLDLWQSAGLNIAPPGGYKDKPNNT 68

RESULT 180
O9LB76 PRELIMINARY; PRT; 89 AA.
AC O9LB76;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

```

```

DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=CHN4952A;
RA Ji X., Telford J.L., Burrone D., Guidotti S., Pagliaccia C.,
RA Rayat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050375; AAF26557.1; -
FT NON_TER 1
FT NON_TER 89
SQ SEQUENCE 89 AA; 9481 MW; 272C6A162EFADB72 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQAGLNIIAPEGGKDKPNT 338
DB 40 IGTLDLWQAGLNIIAPEGGKDKPNT 68

RESULT 181
Q9RE86 PRELIMINARY; PRT; 96 AA.
AC Q9RE86
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PE9022;
RA van Doorn L.J.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390591; CAB64269.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10195 MW; B619D55E8FC320E1 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 VNINKLITASTNVAVKFNINELIVKNG 564
DB 30 VNINKLITASTNVAVKFNINELIVKNG 58

RESULT 182
Q9RE78 PRELIMINARY; PRT; 96 AA.
AC Q9RE78
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PE9022;
RA van Doorn L.J.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390591; CAB64269.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10195 MW; B619D55E8FC320E1 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 VNINKLITASTNVAVKFNINELIVKNG 564
DB 30 VNINKLITASTNVAVKFNINELIVKNG 58

RESULT 183
Q9RE77 PRELIMINARY; PRT; 96 AA.
AC Q9RE77
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JA76;
RA van Doorn L.J.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390609; CAB64284.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10313 MW; 1274D55DF7816A7E CRC64;

Query Match 2.2%; Score 29; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 SVGEYTHFSEDIGSQSRINTVLEGTGRS 594
DB 60 SVGEYTHFSEDIGSQSRINTVLEGTGRS 88

```

```

AC Q9RE78;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PO67;
RA van Doorn L.J.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390603; CAB64279.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10198 MW; 9B19D55DF59E87D0 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 SVGEYTHFSEDIGSQSRINTVLEGTGRS 594
DB 60 SVGEYTHFSEDIGSQSRINTVLEGTGRS 88

```

```

RESULT 183
Q9RE77 PRELIMINARY; PRT; 96 AA.
AC Q9RE77
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JA76;
RA van Doorn L.J.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390609; CAB64284.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10313 MW; 1274D55DF7816A7E CRC64;

Query Match 2.2%; Score 29; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 SVGEYTHFSEDIGSQSRINTVLEGTGRS 594
DB 60 SVGEYTHFSEDIGSQSRINTVLEGTGRS 88

```

```

QY 566 SVGEYTHFSEDIGSQSRINTVRLTGTRS 594
Db 60 SVGEYTHFSEDIGSQSRINTVRLTGTRS 88

RESULT 184
ID Q9LB60 PRELIMINARY; PRT; 96 AA.
AC Q9LB60;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PO32, AND PO30;
RA van Doorn L.J.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PO32, AND PO30;
RX MEDLINE=98371099; PubMed=9705399;
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,
RA Atherton J.C., Blaser M.J., Quint W.G.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390594; CAB64272.1; -
DR EMBL; AJ390593; CAB64271.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10210 MW; EC19D559F8F26282 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 566 SVGEYTHFSEDIGSQSRINTVRLTGTRS 594
Db 60 SVGEYTHFSEDIGSQSRINTVRLTGTRS 88

RESULT 185
ID Q9LB60 PRELIMINARY; PRT; 96 AA.
AC Q9LB60;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHN5147B;
RA Ji X., Telford J.L., Burrioni D., Guidotti S., Pagliaccia C.,
RA Rayat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050392; AAF26574.1; -
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10160 MW; 9AD3FF498BE37CDA CRC64;

```

```

Query Match 2.2%; Score 29; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQSAGLNIIAPPEGYKDKPNNT 338
Db 40 IGTLDLWQSAGLNIIAPPEGYKDKPNNT 68

RESULT 186
ID Q9RE85 PRELIMINARY; PRT; 97 AA.
AC Q9RE85;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VACA PROTEIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TH9022;
RA van Doorn L.J.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TH9022;
RX MEDLINE=98371099; PubMed=9705399;
RA van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K.,
RA Atherton J.C., Blaser M.J., Quint W.G.;
RT "Expanding allelic diversity of Helicobacter pylori vacA.";
RL J. Clin. Microbiol. 36:2597-2603(1998).
DR EMBL; AJ390592; CAB64270.1; -
FT NON_TER 1
FT NON_TER 97
SQ SEQUENCE 97 AA; 10473 MW; D088F8FC8B83A45 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 566 SVGEYTHFSEDIGSQSRINTVRLTGTRS 594
Db 61 SVGEYTHFSEDIGSQSRINTVRLTGTRS 89

RESULT 187
ID Q9LB63 PRELIMINARY; PRT; 98 AA.
AC Q9LB63;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHN5147A;
RA Ji X., Telford J.L., Burrioni D., Guidotti S., Pagliaccia C.,
RA Rayat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050389; AAF26571.1; -
FT NON_TER 1

```

```
FT NON_TER 98
SQ SEQUENCE 98 AA; 10402 MW; 2F5A4170D4ED1E9F CRC64;

Query Match 2.2%; Score 29; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches: 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 338
DB 35 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 63

RESULT 188
Q9K331 PRELIMINARY; PRT; 118 AA.
AC Q9K331;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN5140B, AND CHN5027A;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050417; AAF26599.1; -
DR EMBL; AF050380; AAF26562.1; -
FT NON_TER 118
FT NON_TER 118
SQ SEQUENCE 118 AA; 12472 MW; 508C329875C2033E CRC64;

Query Match 2.2%; Score 29; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.7e-21;
Matches: 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 338
DB 40 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 68

RESULT 189
Q9LB78 PRELIMINARY; PRT; 120 AA.
AC Q9LB78;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN4901A;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050373; AAF26555.1; -
FT NON_TER 120
FT NON_TER 120
SQ SEQUENCE 120 AA; 12757 MW; AB6C14EC79FA3C16 CRC64;

FT NON_TER 98
SQ SEQUENCE 98 AA; 10402 MW; 2F5A4170D4ED1E9F CRC64;

Query Match 2.2%; Score 29; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches: 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 338
DB 35 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 63

RESULT 188
Q9K331 PRELIMINARY; PRT; 118 AA.
AC Q9K331;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN5140B, AND CHN5027A;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050417; AAF26599.1; -
DR EMBL; AF050380; AAF26562.1; -
FT NON_TER 118
FT NON_TER 118
SQ SEQUENCE 118 AA; 12472 MW; 508C329875C2033E CRC64;

Query Match 2.2%; Score 29; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.7e-21;
Matches: 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 338
DB 40 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 68

RESULT 189
Q9LB78 PRELIMINARY; PRT; 120 AA.
AC Q9LB78;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN4901A;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050373; AAF26555.1; -
FT NON_TER 120
FT NON_TER 120
SQ SEQUENCE 120 AA; 12757 MW; AB6C14EC79FA3C16 CRC64;
```

```
Query Match 2.2%; Score 29; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.7e-21;
Matches: 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 338
DB 40 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 68

RESULT 190
Q9LBA2 PRELIMINARY; PRT; 123 AA.
AC Q9LBA2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN3503C;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050348; AAF26530.1; -
FT NON_TER 123
FT NON_TER 123
SQ SEQUENCE 123 AA; 13002 MW; 1222E0A5E0FBEB9F CRC64;

Query Match 2.2%; Score 29; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches: 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 338
DB 40 IGTLDLWQSAGLNIIAPPEGGYKDKPNT 68

RESULT 191
Q9K2T4 PRELIMINARY; PRT; 136 AA.
AC Q9K2T4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHN4611A, AND CHN3295F;
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,
RA Rayrat J.M., Xu G., Rappuoli R.;
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050370; AAF26552.1; -
DR EMBL; AF050338; AAF26520.1; -
FT NON_TER 136
FT NON_TER 136
SQ SEQUENCE 136 AA; 14332 MW; A474FBE0C172E0F6 CRC64;
```

Query Match 2.2%; Score 29; DB 2; Length 136;  
Best Local Similarity 100.0%; Pred. No. 1.9e-21;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDMQSLNIIAPPEGKYDKPNT 338  
DB 40 IGTLDMQSLNIIAPPEGKYDKPNT 68

RESULT 192  
Q9LBC4 PRELIMINARY; PRT; 139 AA.  
ID Q9LBC4  
AC Q9LBC4  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CHN3132B;  
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050325; AAF26507.1;  
FT NON\_TER 1  
FT NON\_TER 139  
SQ SEQUENCE 139 AA; 14719 MW; 5B082DF384DEF8A4 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 139;  
Best Local Similarity 100.0%; Pred. No. 2e-21;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDMQSLNIIAPPEGKYDKPNT 338  
DB 40 IGTLDMQSLNIIAPPEGKYDKPNT 68

RESULT 193  
Q9LB98 PRELIMINARY; PRT; 143 AA.  
ID Q9LB98  
AC Q9LB98  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CHN3554A;  
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050352; AAF26534.1;  
FT NON\_TER 1  
FT NON\_TER 143  
SQ SEQUENCE 143 AA; 15070 MW; 3CC67B7A9C72768 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 143;  
Best Local Similarity 100.0%; Pred. No. 2e-21;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDMQSLNIIAPPEGKYDKPNT 338  
DB 40 IGTLDMQSLNIIAPPEGKYDKPNT 68

RESULT 194  
Q9LBB8 PRELIMINARY; PRT; 149 AA.  
ID Q9LBB8  
AC Q9LBB8  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CHN1811H;  
RA Ji X., Telford J.L., Burroni D., Guidotti S., Pagliaccia C.,  
RA Rayrat J.M., Xu G., Rappuoli R.;  
RT "Allelic variation of vacA gene in the Chinese Helicobacter pylori.";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050331; AAF26513.1;  
FT NON\_TER 1  
FT NON\_TER 149  
SQ SEQUENCE 149 AA; 15655 MW; 77209FDBAD742ECA CRC64;

Query Match 2.2%; Score 29; DB 2; Length 149;  
Best Local Similarity 100.0%; Pred. No. 2.1e-21;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 IGTLDMQSLNIIAPPEGKYDKPNT 338  
DB 40 IGTLDMQSLNIIAPPEGKYDKPNT 68

RESULT 195  
Q9LAJ6 PRELIMINARY; PRT; 214 AA.  
ID Q9LAJ6  
AC Q9LAJ6  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE VACUOLATING CYTOTOXIN A (FRAGMENT).  
GN VACA.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-V290;  
RA Pan S.K., Kuo C.H., Su I.J., Su R., Wang W.C.;  
RT "Characterization of Helicobacter pylori strains isolated from  
RT Helicobacter pylori-infected spouses in Taiwan.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF132303; AAF61326.1;  
FT NON\_TER 1  
FT NON\_TER 214  
SQ SEQUENCE 214 AA; 23662 MW; F1F7F8BF574A2214 CRC64;

Query Match 2.2%; Score 29; DB 2; Length 214;  
Best Local Similarity 100.0%; Pred. No. 2.9e-21;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 NYFDARNIKNVEITNKLAFGPGSPWGTS 647  
DB 40 NYFDARNIKNVEITNKLAFGPGSPWGTS 647

Db 73 NYDARNIKNVEITNKLAFGQSPWCTS 101

```
RESULT 196
Q9R845 ID O87750 PRELIMINARY; PRT; 148 AA.
AC O87750;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-JUN-2000 (TREMELrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC53;
RX MEDLINE=98445420; PubMed=9770535;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009434; CAA08709.1; -
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16633 MW; 06693562307A680E CRC64;
```

```
Query Match 2.2%; Score 28; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 NFTGGDLVNMOKATRLRGQFNGNSFTS 177
|||||
Db 82 NFTGGDLVNMOKATRLRGQFNGNSFTS 109
```

```
RESULT 197
Q9R845 ID Q9R845 PRELIMINARY; PRT; 148 AA.
AC Q9R845;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMELrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC58;
RX MEDLINE=98445420; PubMed=9770535;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009436; CAA08711.1; -
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16632 MW; 0669356890D0680E CRC64;
```

```
Query Match 2.2%; Score 28; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 NFTGGDLVNMOKATRLRGQFNGNSFTS 177
|||||
Db 82 NFTGGDLVNMOKATRLRGQFNGNSFTS 109
```

```
RESULT 198
Q9R841 ID Q9R841 PRELIMINARY; PRT; 148 AA.
AC Q9R841;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMELrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FA6;
RX MEDLINE=98445420; PubMed=9770535;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009440; CAA08715.1; -
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16633 MW; 066EFB2FF9506F6E CRC64;
```

```
Query Match 2.2%; Score 28; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 NFTGGDLVNMOKATRLRGQFNGNSFTS 177
|||||
Db 82 NFTGGDLVNMOKATRLRGQFNGNSFTS 109
```

```
RESULT 199
Q9R840 ID Q9R840 PRELIMINARY; PRT; 148 AA.
AC Q9R840;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMELrel. 14, Last annotation update)
DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FB2;
RX MEDLINE=98445420; PubMed=9770535;
RA Suerbaum S., Maynard Smith J., Bapumia K., Morelli G., Smith N.H.,
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009441; CAA08716.1; -
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16602 MW; 06673B9ECDDBA80E CRC64;
```

```
Query Match 2.2%; Score 28; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 NFTGGDLVNMOKATRLRGQFNGNSFTS 177
|||||
Db 82 NFTGGDLVNMOKATRLRGQFNGNSFTS 109
```

Wed Aug 29 10:20:39 2001

```

RESULT 200
Q9R839 PRELIMINARY; PRT; 148 AA.
ID AC Q9R839;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE DE VACUOLATING CYTOTOXIN (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FB3;
RX MEDLINE=9845420; PubMed=9770535; Bapumia K., Morelli G., Smith N.H.,
RA Suerbaum S., Maynard Smith J., Achtman M.;
RA Kunstmann E., Dyrek I., Achtman M.;
RT "Free recombination within Helicobacter pylori.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12619-12624(1998).
DR EMBL; AJ009442; CAA08717.1; -.
FT NON_TER 1
FT NON_TER 148
FT NON_TER 148
SQ SEQUENCE 148 AA; 16633 MW; FD27D0055860D30E CRC64;
```

```

Query Match 2.2%; Score 28; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. NO. 2.3e-20; Mismatches 0; Gaps 0;
Matches 28; Conservative 0; Indels 0;
```

```

QY 150 NFTGGDLVNMQRKATLRLGQFNGNSFTS 177
DB 82 NFTGGDLVNMQRKATLRLGQFNGNSFTS 109
```

Search completed: August 29, 2001, 09:38:05  
Job time: 226 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2001, 09:33:09 ; Search time: 24.11 Seconds  
(without alignments)

4094.659 Million cell updates/sec

Title: US-09-360-934A-3

Perfect score: 1296

Sequence: 1 MEIQTHRKINRPLVSLAV.....HNLISNIGHFASNLGMRYSF 1296

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 200 summaries

Database :

PIR 68: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	8.5	1288	2	E71884
2	99	7.6	1287	2	B53739
3	87	6.7	1290	2	G64630
4	87	6.7	1291	2	S44983
5	20	1.5	513	2	A53739
6	8	0.6	164	2	T06913
7	8	0.6	334	2	C84935
8	8	0.6	346	1	I40372
9	8	0.6	611	2	S61147
10	8	0.6	637	2	B81978
11	8	0.6	637	2	D81034
12	8	0.6	935	2	T19011
13	8	0.6	1140	2	B70729
14	8	0.6	2893	2	A64556
15	8	0.6	2902	2	C71953
16	8	0.6	3603	1	D69681
17	7	0.5	22	2	PH1678
18	7	0.5	22	2	PH1679
19	7	0.5	23	2	PH1681
20	7	0.5	23	2	PH1682
21	7	0.5	23	2	PH1684
22	7	0.5	24	2	PH1683
23	7	0.5	24	2	PH1685
24	7	0.5	25	2	PH1686
25	7	0.5	25	2	PH1700
26	7	0.5	34	2	A61491
27	7	0.5	84	2	PL0127
28	7	0.5	90	2	I76722
29	7	0.5	90	2	H85984

30	7	0.5	99	2	A70366
31	7	0.5	103	2	S72589
32	7	0.5	121	1	B2AG55
33	7	0.5	121	1	B2AGAG
34	7	0.5	127	2	E75555
35	7	0.5	128	2	T41487
36	7	0.5	141	2	C83475
37	7	0.5	151	2	PQ0506
38	7	0.5	159	2	S33585
39	7	0.5	161	2	JQ1693
40	7	0.5	161	2	S65777
41	7	0.5	161	2	H84518
42	7	0.5	161	2	T45148
43	7	0.5	162	2	T08154
44	7	0.5	162	2	F69017
45	7	0.5	164	2	T31280
46	7	0.5	177	2	D83111
47	7	0.5	177	2	S39859
48	7	0.5	182	2	E64678
49	7	0.5	182	2	C71839
50	7	0.5	188	2	F82285
51	7	0.5	206	2	S08632
52	7	0.5	207	1	D64601
53	7	0.5	207	2	C71914
54	7	0.5	215	2	S78275
55	7	0.5	215	2	F75166
56	7	0.5	225	2	T17795
57	7	0.5	226	2	A53273
58	7	0.5	227	2	I54426
59	7	0.5	227	2	G69491
60	7	0.5	229	2	A48381
61	7	0.5	229	2	I55971
62	7	0.5	229	2	T27840
63	7	0.5	235	2	H69488
64	7	0.5	236	2	I67432
65	7	0.5	236	2	I53744
66	7	0.5	236	2	JC7383
67	7	0.5	239	1	TVHUA1
68	7	0.5	239	2	D72747
69	7	0.5	247	2	S37342
70	7	0.5	250	2	A71684
71	7	0.5	252	2	A46505
72	7	0.5	252	2	A83089
73	7	0.5	252	2	C83778
74	7	0.5	253	2	S15684
75	7	0.5	253	2	JC2388
76	7	0.5	254	1	HLHDA
77	7	0.5	255	1	HLMSD
78	7	0.5	255	1	HLMSEA
79	7	0.5	255	2	S06316
80	7	0.5	255	2	A45881
81	7	0.5	257	2	S28473
82	7	0.5	259	2	A38284
83	7	0.5	259	2	JS0660
84	7	0.5	260	2	D86722
85	7	0.5	262	2	D82918
86	7	0.5	263	2	C70881
87	7	0.5	263	2	A31299
88	7	0.5	264	2	E82346
89	7	0.5	269	2	H71727
90	7	0.5	283	1	LNPHLS
91	7	0.5	286	2	A84980
92	7	0.5	291	2	A83334
93	7	0.5	293	2	E86474
94	7	0.5	294	2	B64649
95	7	0.5	294	2	D71938
96	7	0.5	295	2	A60131
97	7	0.5	298	2	A71425
98	7	0.5	299	2	G64541
99	7	0.5	299	2	B71967
100	7	0.5	299	2	A96665
101	7	0.5	302	1	JN0794
102	7	0.5	311	2	T33899

hypothetical prote  
hypothetical prote  
virB2 protein - Ag  
conserved hypothet  
very hypothetical  
hypothetical prote  
hypothetical prote  
lipoprotein smpA p  
pathogenesis-relat  
pathogenesis-relat  
pathogenesis-relat  
methyl coenzyme M  
methyl coenzyme M  
benzoate 1,2-dioxy  
transcription anti  
transcription anti  
NADH dehydrogenase  
NADH dehydrogenase  
conserved hypothet  
nodulin-21 - soybe  
phosphoserine phos  
phosphoserine phos  
ribosomal protein  
hypothetical prote  
hypothetical prote  
MHC class II histo  
MHC H2-w28-E alpha  
conserved hypothet  
MHC class II histo  
MHC H2-IE-alpha ce  
hypothetical prote  
SSU ribosomal prot  
BCL-2 - rat (fragm  
gene bcl-2 protein  
B-cell lymphoma 2  
transforming prote  
probable proteasom  
chitinase (EC 3.2.  
probable conjugal  
SLA-DRad (MHC Clas  
conserved hypothet  
ferrichrome ABC tr  
MHC class II histo  
class II histocomp  
H-2 class II histo  
H-2 class II histo  
class II histocomp  
MHC class II histo  
MHC class II histo  
rfbH protein - Vib  
homeotic protein Q  
homeotic protein H  
hypothetical prote  
conserved hypothet  
probable thya prot  
chymotrypsin (EC 3  
lipopolysaccharide  
DNA polymerase III  
lectin precursor -  
DNA-directed DNA p  
hypothetical prote  
unknown protein (i  
ATP-binding protei  
hypothetical prote  
homeotic protein X  
hypothetical prote  
cell binding facto  
probable peptidyl-  
protein F22C12.24  
phosphate butyryl  
hypothetical prote

103	7	0.5	313	2	A83951	103	7	0.5	176	7	0.5	484	2	T34504	hypothetical prote
104	7	0.5	318	2	B84291	104	7	0.5	177	7	0.5	488	2	T32149	hypothetical prote
105	7	0.5	319	1	A43452	105	7	0.5	178	7	0.5	492	2	T28025	hypothetical prote
106	7	0.5	319	2	C85503	106	7	0.5	179	7	0.5	493	2	F96696	protein FN21.12 [
107	7	0.5	319	2	F82100	107	7	0.5	180	7	0.5	495	2	T42758	double-stranded RN
108	7	0.5	322	2	S73795	108	7	0.5	181	7	0.5	497	2	T47715	hypothetical prote
109	7	0.5	325	2	G69580	109	7	0.5	182	7	0.5	498	1	S43833	glyceraldehyde-3-p
110	7	0.5	325	2	E65084	110	7	0.5	183	7	0.5	502	2	S52648	myo-inositol-1-pho
111	7	0.5	328	2	B84263	111	7	0.5	184	7	0.5	507	2	T36370	probable sensory h
112	7	0.5	329	2	A83405	112	7	0.5	185	7	0.5	509	2	T08436	myo-inositol-1-pho
113	7	0.5	330	1	A26478	113	7	0.5	186	7	0.5	510	2	T04399	myo-inositol-1-pho
114	7	0.5	330	2	B65179	114	7	0.5	187	7	0.5	510	2	T01647	myo-inositol-1-pho
115	7	0.5	330	2	H86061	115	7	0.5	188	7	0.5	510	2	S60302	myo-inositol-1-pho
116	7	0.5	331	2	S74931	116	7	0.5	189	7	0.5	510	2	D84610	probable myo-inosi
117	7	0.5	331	2	G96750	117	7	0.5	190	7	0.5	510	2	T50021	myo-inositol-1-pho
118	7	0.5	333	2	D83585	118	7	0.5	191	7	0.5	510	2	T39930	replication protei
119	7	0.5	334	1	RDECEP	119	7	0.5	192	7	0.5	511	2	T05017	myo-inositol-1-pho
120	7	0.5	334	2	D86087	120	7	0.5	193	7	0.5	511	2	T10964	myo-inositol-1-pho
121	7	0.5	335	2	D83582	121	7	0.5	194	7	0.5	512	2	T12438	myo-inositol-1-pho
122	7	0.5	340	2	T51045	122	7	0.5	195	7	0.5	514	2	E82214	galactoside ABC tr
123	7	0.5	342	2	S57510	123	7	0.5	196	7	0.5	515	2	A69759	1-pyrroline-5-carb
124	7	0.5	344	2	C64978	124	7	0.5	197	7	0.5	515	2	A83992	1-pyrroline-5-carb
125	7	0.5	344	2	A44164	125	7	0.5	198	7	0.5	515	2	D84142	hypothetical prote
126	7	0.5	345	2	S72490	126	7	0.5	199	7	0.5	519	2	G86335	hypothetical prote
127	7	0.5	345	2	E71843	127	7	0.5	200	7	0.5	519	2	S38921	hypothetical prote
128	7	0.5	346	2	C82052	128	7	0.5							
129	7	0.5	351	2	G71372	129	7	0.5							
130	7	0.5	351	2	C81901	130	7	0.5							
131	7	0.5	352	1	B64673	131	7	0.5							
132	7	0.5	353	1	E64581	132	7	0.5							
133	7	0.5	355	2	C39725	133	7	0.5							
134	7	0.5	356	2	G84904	134	7	0.5							
135	7	0.5	358	2	T45934	135	7	0.5							
136	7	0.5	365	2	E84890	136	7	0.5							
137	7	0.5	367	2	T16913	137	7	0.5							
138	7	0.5	370	2	D70076	138	7	0.5							
139	7	0.5	376	2	S73941	139	7	0.5							
140	7	0.5	383	2	T52651	140	7	0.5							
141	7	0.5	388	2	E64707	141	7	0.5							
142	7	0.5	388	2	C71813	142	7	0.5							
143	7	0.5	389	2	G70810	143	7	0.5							
144	7	0.5	390	2	I39570	144	7	0.5							
145	7	0.5	391	2	T30149	145	7	0.5							
146	7	0.5	393	2	T29412	146	7	0.5							
147	7	0.5	395	2	B64709	147	7	0.5							
148	7	0.5	395	2	G71811	148	7	0.5							
149	7	0.5	396	2	D69378	149	7	0.5							
150	7	0.5	397	2	T35713	150	7	0.5							
151	7	0.5	401	2	T00658	151	7	0.5							
152	7	0.5	403	1	VCBBBL	152	7	0.5							
153	7	0.5	410	2	S40976	153	7	0.5							
154	7	0.5	410	2	B36670	154	7	0.5							
155	7	0.5	410	2	T13255	155	7	0.5							
156	7	0.5	415	1	A34170	156	7	0.5							
157	7	0.5	421	1	S11674	157	7	0.5							
158	7	0.5	422	1	A84237	158	7	0.5							
159	7	0.5	429	2	C83025	159	7	0.5							
160	7	0.5	430	2	T19722	160	7	0.5							
161	7	0.5	430	2	T46317	161	7	0.5							
162	7	0.5	431	2	S47538	162	7	0.5							
163	7	0.5	435	2	E70586	163	7	0.5							
164	7	0.5	437	2	T16477	164	7	0.5							
165	7	0.5	438	2	B83295	165	7	0.5							
166	7	0.5	441	2	T18500	166	7	0.5							
167	7	0.5	448	2	A70398	167	7	0.5							
168	7	0.5	449	2	S67819	168	7	0.5							
169	7	0.5	450	2	E71044	169	7	0.5							
170	7	0.5	451	2	D75170	170	7	0.5							
171	7	0.5	457	2	E69056	171	7	0.5							
172	7	0.5	462	2	T17480	172	7	0.5							
173	7	0.5	463	2	S72992	173	7	0.5							
174	7	0.5	465	2	A82211	174	7	0.5							
175	7	0.5	467	2	E82567	175	7	0.5							

## ALIGNMENTS

## RESULT 1

E71884  
 vacuolating cytotoxin - Helicobacter pylori (strain J99)  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
 C:Accession: E71884  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, G.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.

Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric  
 A:Reference number: A71800; MUID:99120557  
 A:Accession: E71884  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-1288 <ARN>  
 A:Cross-references: GB:AE001511; GB:AE001439; NID:g4155382; PIDN:AAD06400.1; PID:g4

A:Experimental source: strain J99  
 A:Genetics:  
 A:Gene: vacA

Query Match  
 Best Local Similarity 100.0%; Pred. No. 2,8e-106;  
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 DTKNGTATFNNDISLGRFVNLKVDAHTANFKGIDTNGGNTLDFSGVTDKVNINKLITA 544  
 DB 478 DTKNGTATFNNDISLGRFVNLKVDAHTANFKGIDTNGGNTLDFSGVTDKVNINKLITA 537

QY 545 STNVAVKFNINELIVKTNISVGVEYTHFSESDIGSOSRINTVRLTGTRS 594  
 DB 538 STNVAVKFNINELIVKTNISVGVEYTHFSESDIGSOSRINTVRLTGTRS 587

RESULT 2  
 B53739  
 vacuolating cytotoxin precursor - Helicobacter pylori (strain ATCC 49503)

C:Species: Helicobacter pylori  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: B53739; A38137

R:Cover, T.L.; Tummuru, M.K.R.; Cao, P.; Thompson, S.A.; Blaser, M.J.  
 J. Biol. Chem. 269, 10566-10573, 1994

A:Title: Divergence of genetic sequences for the vacuolating cytotoxin among *Helicobacter pylori* (isolate 185-44)  
 A:Reference number: A53739; MUID:94193753  
 A:Accession: B53739  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1287 <COV>

A:Cross-references: GB:U05676; NID:9471727; PIDN:AAAI7657.1; PID:9471729  
 A:Note: parts of this sequence, including the amino end of the mature protein, were determined by J. Biol. Chem. 267, 10570-10575, 1992  
 R:Cover, T.L.; Blaser, M.J.

A:Title: Purification and characterization of the vacuolating toxin from *Helicobacter pylori*  
 A:Reference number: A38137; MUID:92268100  
 A:Accession: A38137  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 34-56 <CO2>

A:Note: sequence extracted from NCBI backbone (NCBIP:103729)  
 C:Genetics:  
 A:Gene: vacA  
 C:Keywords: cytotoxin  
 F:1-33/Domain: signal sequence #status predicted <SIG>

Query Match 7.6%; Score 99; DB 2; Length 1287;  
 Best Local Similarity 100.0%; Pred. No. 1e-94;  
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 VGNAAAMFNNDISATGYKPLIKNSAQDLKNTHEVLLKAKIIGYGVSTGTNGISN 756  
 |||||  
 Db 689 VGNAAAMFNNDISATGYKPLIKNSAQDLKNTHEVLLKAKIIGYGVSTGTNGISN 748

QY 757 VNLEQFKERLALYNNNRMDTCVVRNTDDIKACGMAIG 795  
 |||||  
 Db 749 VNLEQFKERLALYNNNRMDTCVVRNTDDIKACGMAIG 787

RESULT 3  
 G64630  
 vacuolating cytotoxin precursor - *Helicobacter pylori* (strain 26695)  
 C:Species: *Helicobacter pylori*  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 C:Accession: G64630  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne-son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
 A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: G64630  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1290 <TOM>

A:Cross-references: GB:AE000598; GB:AE000511; NID:92314019; PIDN:AAD07935.1; PID:9231402  
 F:1-33/Domain: signal sequence #status predicted <SIG>  
 F:34-1290/Product: vacuolating cytotoxin #status predicted <MAT>

Query Match 6.7%; Score 87; DB 2; Length 1290;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-82;  
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTGNTVDPGLRVNNQVGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNK 506  
 |||||  
 Db 442 VENLTGNTVDPGLRVNNQVGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNK 501

QY 507 VDAHTANFKGIDTNGGNTLDFSGVT 533  
 |||||  
 Db 502 VDAHTANFKGIDTNGGNTLDFSGVT 528

RESULT 4  
 S44983  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

vacuolating cytotoxin precursor - *Helicobacter pylori* (isolate 185-44)  
 C:Species: *Helicobacter pylori*  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: S44983; S44102  
 R:Schmitt, W.; Haas, R.  
 Mol. Microbiol. 12, 307-319, 1994

A:Title: Genetic analysis of the *Helicobacter pylori* vacuolating cytotoxin: structure  
 A:Reference number: S44983; MUID:94335650  
 A:Accession: S44983  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <SCH>

A:Cross-references: EMBL:Z26883; NID:9472941; PIDN:CAA81528.1; PID:9472942  
 C:Keywords: cytotoxin  
 F:1-33/Domain: signal sequence #status predicted <SIG>  
 F:34-1291/Product: vacuolating cytotoxin #status predicted <MAT>

Query Match 6.7%; Score 87; DB 2; Length 1291;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-82;  
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTGNTVDPGLRVNNQVGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNK 506  
 |||||  
 Db 442 VENLTGNTVDPGLRVNNQVGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNK 501

QY 507 VDAHTANFKGIDTNGGNTLDFSGVT 533  
 |||||  
 Db 502 VDAHTANFKGIDTNGGNTLDFSGVT 528

RESULT 5  
 A53739  
 hypothetical vacuolating cytotoxin - *Helicobacter pylori* (strain ATCC 53726) (fragment)  
 C:Species: *Helicobacter pylori*  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: A53739  
 R:Cover, T.L.; Tummuru, M.K.R.; Cao, P.; Thompson, S.A.; Blaser, M.J.  
 J. Biol. Chem. 269, 10566-10573, 1994

A:Title: Divergence of genetic sequences for the vacuolating cytotoxin among *Helicobacter pylori*  
 A:Reference number: A53739; MUID:94193753  
 A:Accession: A53739  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-513 <COV>

A:Cross-references: GB:U05677; NID:9471730; PIDN:AAAI7658.1; PID:9471731  
 A:Note: this strain, designated 87-203, ATCC 53726 tox-, does not possess toxin activ  
 C:Genetics:  
 A:Gene: vacA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

Query Match 1.5%; Score 20; DB 2; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 OPTQVIDGPFAGGKDTVVNI 400  
 |||||  
 Db 44 OPTQVIDGPFAGGKDTVVNI 63

RESULT 6  
 T06913  
 H<sup>+</sup>-transporting ATP synthase (EC 3.6.1.34) chain b' - *Cyanophora paradoxa* cyanelle  
 C:Species: *Cyanophora paradoxa*  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T06913  
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
 A:Reference number: Z15840  
 A:Accession: T06913  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

A:Residues: 1-164 <ST1>  
 A:Cross-references: EMBL:U30821; NID:g1016083; PIDN:AAA81256.1; PID:g1016169  
 A:Experimental source: strain Pringsheim LB555  
 C:Genetics:

A:Gene: atpg  
 A:Genome: cyanelle  
 C:Superfamily: H<sup>+</sup>-transporting ATP synthase chain I  
 C:Keywords: ATP biosynthesis; cyanelle; hydrolase; thylakoid

Query Match 0.6%; Score 8; DB 2; Length 164;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 715 FYKPLIKI 722

Db 46 FYKPLIKI 53

RESULT 7

C84935 N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) [Imported] - Buchnera sp. (str

C:Species: Buchnera sp.

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: C84935

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A

A:Reference number: A84930; MUID:20445173

A:Accession: C84935

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144

A:Experimental source: strain APS

C:Genetics:

A:Gene: argC; BU048

C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase

C:Keywords: oxidoreductase

Query Match 0.6%; Score 8; DB 2; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 209 SGAGRKAS 216

Db 187 SGAGRKAS 194

RESULT 8

I40372 N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) - Bacillus subtilis

N:Alternate names: acetylglutamate semialdehyde dehydrogenase; N-acetylglutamate-gamma-s

C:Species: Bacillus subtilis

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000

C:Accession: I40372; S13592; A26536; S38428; F69588; S20023

R:O'Reilly, M.; Devine, K.M.

Microbiology 140, 1023-1025, 1994

A:Title: Sequence and analysis of the citrulline biosynthetic operon argC-F from Bacillu

A:Reference number: I40372; MUID:94297722

A:Accession: I40372

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-346 <RES>

A:Cross-references: EMBL:Z26919; NID:g408113; PID:g575620

R:Smith, M.C.M.; Mountain, A.; Baumberg, S.

Nucleic Acids Res. 18, 4595, 1990

A:Title: Nucleotide sequence of the Bacillus subtilis argC gene encoding N-acetylglutami

A:Reference number: S12592; MUID:90356403

A:Accession: S12592

A:Molecule type: DNA

A:Residues: 1-340,342-346 <SM1>

A:Cross-references: EMBL:X52834; NID:g39806; PIDN:CAA37016.1; PID:g580828  
 R:Smith, M.C.M.; Mountain, A.; Baumberg, S.  
 Gene 49, 53-60, 1986

A:Title: Sequence analysis of the Bacillus subtilis argC promoter region.

A:Reference number: A26526; MUID:87192000

A:Accession: A26526

A:Molecule type: DNA

A:Residues: 1-56 <SM2>

A:Cross-references: GB:M15420; NID:g142533; PIDN:AAA22448.1; PID:g142534

R:O'Reilly, M.; Devine, K.M.

submitted to the EMBL Data Library, October 1993

A:Reference number: S38428

A:Accession: S38428

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-346 <OAR>

A:Cross-references: EMBL:Z26919

R:Czaplewski, L.G.; North, A.K.; Smith, M.C.M.; Baumberg, S.; Stockley, P.G.

Mol. Microbiol. 6, 267-275, 1992

A:Title: Purification and initial characterization of AhrC: the regulator of arginin

A:Reference number: S20023; MUID:92186717

A:Contents: annotation

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be

C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.

A.; Ehrlich, S.D.; Emmeron, P.T.; Eutian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio,

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardin

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scan

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; S

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshid

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil

A:Reference number: A69580; MUID:96044033

A:Accession: F69588

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-234,'V',236-346 <KUN>

A:Cross-references: GB:299109; GB:299110; GB:AL009126; NID:g2633472; PIDN:CAB12976.1

A:Experimental source: strain 168

C:Genetics:

A:Gene: argC

A:Start codon: TTG

C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase

C:Keywords: oxidoreductase

Query Match 0.6%; Score 8; DB 1; Length 346;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 209 SGAGRKAS 216

Db 181 SGAGRKAS 188

RESULT 9

S61147

TCM10 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein D9476.9; protein YDR350C

C:Species: Saccharomyces cerevisiae

C:Date: 23-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 23-Mar-2001

C:Accession: S61147; S59675

R:Du, Z.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of S. cerevisiae cosmid 9476.

A:Reference number: S61148

A:Accession: S61147

A:Molecule type: DNA

A:Residues: 1-611 <DUZ>  
 A:Cross-references: EMBL:U28372; NID:g849170; PID:g849179; MIPS:YDR350C  
 R:Zhang, Y.; Robinson, K.M.; Lemire, B.D.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: S59675  
 A:Accession: S59675  
 A:Molecule type: DNA  
 A:Residues: 1-591, 'LFGSGEIRHMAIQLIKDQGFPPKFNFDLTLVELVNNIKEPTDSTLF', 'T  
 A:Cross-references: EMBL:U32306; NID:g929984; PID:g929985  
 A:Experimental source: strain MH125  
 C:Genetics:  
 A:Gene: SGD:TCM10  
 A:Cross-references: SGD:S0002758; MIPS:YDR350C  
 A:Map position: 4R

Query Match 0.6%; Score 8; DB 2; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 LYRSLSS 107

Db 452 LYRSLSS 459

RESULT 10

B81978  
 probable 1-deoxyxylulose-5-phosphate synthase NMA0589 [Imported] - Neisseria meningitidis  
 C:Species: Neisseria meningitidis  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
 C:Accession: B81978  
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
 ; Holroyd, S.; Jagels, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream,  
 Nature 404, 502-506, 2000  
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
 A:Reference number: B81775; MUID:2022556  
 A:Accession: B81978  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-637 <PAR>  
 A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83880.1; PID:g737932  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: dxs; NMA0589  
 C:Superfamily: hypothetical protein C2814

Query Match 0.6%; Score 8; DB 2; Length 637;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1284 GHFASNLG 1291

Db 44 GHFASNLG 51

RESULT 11

D81034  
 1-deoxyxylulose-5-phosphate synthase NMB1867 [Imported] - Neisseria meningitidis (strain  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C:Accession: D81034  
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: A81000; MUID:20175755  
 A:Accession: D81034  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-637 <TET>

A:Cross-references: GB:AE002536; GB:AE002098; NID:g7227115; PIDN:AAF42201.1; PID:g722  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB1867  
 C:Superfamily: hypothetical protein C2814

Query Match 0.6%; Score 8; DB 2; Length 637;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1284 GHFASNLG 1291

Db 44 GHFASNLG 51

RESULT 12

T19011  
 hypothetical protein C06C6.7 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T19011  
 R:McMurray, A.  
 submitted to the EMBL Data Library, March 1997  
 A:Reference number: Z19059  
 A:Accession: T19011  
 A>Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-935 <WIL>  
 A:Cross-references: EMBL:Z93374; PIDN:CAB07557.1; GSPDB:GN00023; CESP:C06C6.7  
 A:Experimental source: clone C06C6  
 C:Genetics:  
 A:Gene: CESP:C06C6.7  
 A:Map position: 5  
 A:Introns: 28/1; 55/1; 80/1; 801/2; 865/2

Query Match 0.6%; Score 8; DB 2; Length 935;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VGALVSIT 27

Db 609 VGALVSIT 616

RESULT 13

B70729  
 hypothetical protein RV2566 - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C:Accession: B70729  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno.  
 A:Reference number: A70500; MUID:98295987  
 A:Accession: B70729  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1140 <COL>  
 A:Cross-references: GB:Z77250; GB:AL123456; NID:g3261617; PIDN:CAB01049.1; PID:e25533  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: RV2566

Query Match 0.6%; Score 8; DB 2; Length 1140;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 ITVDGDLR 461  
 Db 1046 ITVDGDLR 1053

RESULT 14  
 A64556  
 toxin-like outer membrane protein HP0289 - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 C:Accession: A64556  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:197394467  
 A:Accession: A64556  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2893 <TOM>  
 A:Cross-references: GB:AE000547; GB:AE000511; NID:92313377; PIDN:AAD07355.1; PID:9231338

Query Match 0.6%; Score: 8; DB 2; Length 2893;  
 Best Local Similarity 100.0%; Pred. No. 99;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
 Db 176 NRVGSGAG 183

RESULT 15  
 C71953  
 toxin-like outer membrane protein jhp0274 - Helicobacter pylori (strain J99)  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99  
 C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
 C:Accession: C71953  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.  
 A:Reference number: A71800; MUID:99120557  
 A:Accession: C71953  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2902 <ARN>  
 A:Cross-references: GB:AE001464; GB:AE001439; NID:94154789; PIDN:AAD05855.1; PID:9415479  
 A:Experimental source: strain J99  
 A:Genetics: jhp0274

Query Match 0.6%; Score: 8; DB 2; Length 2902;  
 Best Local Similarity 100.0%; Pred. No. 99;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
 Db 185 NRVGSGAG 192

RESULT 16  
 D69681  
 peptidase synthetase ppsD - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Nov-2000  
 C:Accession: D69681  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bortel, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Giech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lard, A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; M Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Porte Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Sca A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Keuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchi T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshi A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis A:Title: The complete genome sequence of the Gram-positive bacterium A:Reference number: A69580; MUID:98044033  
 A:Accession: D69681  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-3603 <KUN>  
 A:Cross-references: GB:299113; GB:AL00126; NID:92634090; PIDN:CAB13714.1; PID:9263 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: ppsD  
 C:Superfamily: surfactin synthetase; acetate--CoA ligase homology; acyl carrier pro C:Keywords: carrier protein; phosphopantetheine; phosphoprotein  
 F:509-952/Domain: acetate--CoA ligase homology <ACLP1>  
 F:969-1037/Domain: acyl carrier protein homology <ACLP1>  
 F:1540-1983/Domain: acetate--CoA ligase homology <ACLP2>  
 F:2000-2068/Domain: acyl carrier protein homology <ACLP2>  
 F:2579-3019/Domain: acetate--CoA ligase homology <ACLP3>  
 F:3037-3104/Domain: acyl carrier protein homology <ACLP3>  
 F:1001,2032,3069/Binding site: phosphopantetheine (Ser) (covalent) #status predicte

Query Match 0.6%; Score: 8; DB 1; Length 3603;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 LTLOASEG 227  
 Db 2455 LTLOASEG 2462

RESULT 17  
 PH1678  
 Ig heavy chain V region (clone NP-6-9) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1678  
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
 J. Exp. Med. 178, 295-307, 1993  
 A:Title: Antigen-driven B cell differentiation in vivo.  
 A:Reference number: PH1675; MUID:93301607  
 A:Accession: PH1678  
 A:Molecule type: mRNA  
 A:Residues: 1-22 <MCH>  
 A:Experimental source: B cell  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score: 7; DB 2; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYIG 1211  
 Db 12 EARYYIG 18

RESULT 18  
 PH1679  
 Ig heavy chain V region (clone NP-6-10) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)

C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1679  
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
 J. Exp. Med. 178, 295-307, 1993  
 A:Title: Antigen-driven B cell differentiation in vivo.  
 A:Reference number: PH1675; MUID:93301607

A:Accession: PH1679  
 A:Molecule type: mRNA  
 A:Residues: 1-22 <MCH>  
 A:Experimental source: B cell  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 7; DB 2; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYG 1211  
 |||||  
 Db 12 EARYYG 18

RESULT 19  
 PH1681  
 Ig heavy chain V region (clone NP-6-12) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1681  
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
 J. Exp. Med. 178, 295-307, 1993  
 A:Title: Antigen-driven B cell differentiation in vivo.  
 A:Reference number: PH1675; MUID:93301607

A:Accession: PH1681  
 A:Molecule type: mRNA  
 A:Residues: 1-23 <MCH>  
 A:Experimental source: B cell  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 7; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYG 1211  
 |||||  
 Db 12 EARYYG 18

RESULT 20  
 PH1682  
 Ig heavy chain V region (clone NP-6-13) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1682  
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
 J. Exp. Med. 178, 295-307, 1993  
 A:Title: Antigen-driven B cell differentiation in vivo.  
 A:Reference number: PH1675; MUID:93301607

A:Accession: PH1682  
 A:Molecule type: mRNA  
 A:Residues: 1-23 <MCH>  
 A:Experimental source: B cell  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 7; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYG 1211

Db 12 EARYYG 18  
 |||||

RESULT 21  
 PH1694  
 Ig heavy chain V region (clone NP-7-7) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1694  
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
 J. Exp. Med. 178, 295-307, 1993  
 A:Title: Antigen-driven B cell differentiation in vivo.  
 A:Reference number: PH1675; MUID:93301607

A:Accession: PH1694  
 A:Molecule type: mRNA  
 A:Residues: 1-23 <MCH>  
 A:Experimental source: B cell  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 7; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYG 1211  
 |||||  
 Db 12 EARYYG 18

RESULT 22  
 PH1683  
 Ig heavy chain V region (clone NP-6-14) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1683  
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
 J. Exp. Med. 178, 295-307, 1993  
 A:Title: Antigen-driven B cell differentiation in vivo.  
 A:Reference number: PH1675; MUID:93301607

A:Accession: PH1683  
 A:Molecule type: mRNA  
 A:Residues: 1-24 <MCH>  
 A:Experimental source: B cell  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 7; DB 2; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYG 1211  
 |||||  
 Db 12 EARYYG 18

RESULT 23  
 PH1685  
 Ig heavy chain V region (clone NP-6-16) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1685  
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
 J. Exp. Med. 178, 295-307, 1993  
 A:Title: Antigen-driven B cell differentiation in vivo.  
 A:Reference number: PH1675; MUID:93301607

A:Accession: PH1685  
 A:Molecule type: mRNA  
 A:Residues: 1-24 <MCH>  
 A:Experimental source: B cell  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 7; DB 2; Length 24;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYVG 1211  
|||||||  
Db 12 EARYYVG 18

RESULT 24  
PH1686  
Ig heavy chain V region (clone NP-6-17) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
C;Accession: PH1686  
R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
J. Exp. Med. 178, 295-307, 1993  
A;Title: Antigen-driven B cell differentiation in vivo.  
A;Reference number: PH1675; MUID:93301607  
A;Accession: PH1686  
A;Molecule type: mRNA  
A;Residues: 1-25 <MCH>  
A;Experimental source: B cell  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 7; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYVG 1211  
|||||||  
Db 12 EARYYVG 18

RESULT 25  
PH1700  
Ig heavy chain V region (clone NP-7-13) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
C;Accession: PH1700  
R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
J. Exp. Med. 178, 295-307, 1993  
A;Title: Antigen-driven B cell differentiation in vivo.  
A;Reference number: PH1675; MUID:93301607  
A;Accession: PH1700  
A;Molecule type: mRNA  
A;Residues: 1-25 <MCH>  
A;Experimental source: B cell  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 7; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYVG 1211  
|||||||  
Db 12 EARYYVG 18

RESULT 26  
A61491  
seed protein ws-1 - winged bean (fragment)  
C;Species: Psophocarpus tetragonolobus (winged bean)  
C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 17-Mar-2000  
C;Accession: A61491  
R;Hirano, H.

J. Protein Chem. 8, 115-130, 1989  
A;Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two  
A;Reference number: A61491; MUID:89351606  
A;Accession: A61491  
A;Status: preliminary  
A;Molecule type: protein  
C;Residues: 1-34 <HIR>  
C;Superfamily: plant Kunitz-type proteinase inhibitor  
C;Keywords: seed

Query Match 0.5%; Score 7; DB 2; Length 34;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IIPAIVG 46  
|||||||  
Db 22 IIPAIVG 28

RESULT 27  
PL0127  
H-2 class II histocompatibility antigen I-E alpha chain - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 12-Apr-1995  
C;Accession: PL0127  
R;Schiffenbauer, J.; McCarthy, D.M.; Nygard, N.R.; Woulfe, S.L.; Didier, D.K.; Schwach  
J. Exp. Med. 170, 971-984, 1989  
A;Title: A unique sequence of the NZW I-E beta chain and its possible contribution to  
A;Reference number: PL0126; MUID:89361262  
A;Accession: PL0127  
A;Molecule type: mRNA  
A;Residues: 1-84 <SCH>  
A;Experimental source: strain NZW  
A;Note: the sequence shown here is from the second exon; it is identical with haplot  
A;Note: the authors do not show the first codon for Ile  
A;Note: the authors translated the codon CAA for residue 67 as Lys  
C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 0.5%; Score 7; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAGL 1115  
|||||||  
Db 54 FEAQAGL 60

RESULT 28  
I76722  
phosphocarrier protein Npr (nitrogen related hpr) - Escherichia coli  
C;Species: Escherichia coli  
C;Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 20-Aug-1999  
C;Accession: I76722; H65111; S38619  
R;Jones, D.H.A.; Franklin, C.F.H.; Thomas, C.M.  
Microbiology 140, 1035-1043, 1994  
A;Title: Molecular analysis of the operon which encodes the RNA polymerase sigma fac  
A;Reference number: I57054; MUID:94297724  
A;Accession: I76722  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-90 <RES>  
A;Cross-references: EMBL:Z27094; NID:9414884; PIDN:CAA81621.1; PID:9414889  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617  
A;Accession: H65111  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA



A:Residues: 1-90 <BLAT>  
 A:Cross-references: GB:AE000400; GB:U00096; NID:92367203; PIDN:AAC76238.1; PID:91789599;  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:

A:Gene: ptsO  
 C:Superfamily: phosphotransferase system phosphohistidine-containing protein; phosphotransferase system phosphohistidine; phosphoprotein  
 C:Keywords: phosphotransferase system phosphohistidine-containing protein homology <H  
 F:9-86/Domain: phosphotransferase system phosphohistidine-containing protein homology <H  
 F:16/Binding site: phosphate (His) (covalent) (by phosphotransferase system enzyme I) #8  
 F:48/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 0.5%; Score 7; DB 2; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 629 VEITNKL 635

Db 7 VEITNKL 13

RESULT 29

H85984

hypothetical protein ptsO [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001

C:Accession: H85984

R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,  
 Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: H85984

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <STO>

A:Cross-references: GB:AE005174; NID:912517824; PIDN:AAG58340.1; GSPDB:GN00145; UWGP:245  
 A:Experimental source: strain O157:H7, substrain EDU933

C:Genetics:

A:Gene: ptsO

Query Match 0.5%; Score 7; DB 2; Length 90;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 629 VEITNKL 635

Db 7 VEITNKL 13

RESULT 30

A70366

hypothetical protein aq\_754 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C>Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999

C:Accession: A70366

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
 V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666

A:Accession: A70366

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-99 <AQF>

A:Cross-references: GB:AE000706; NID:929833327; PIDN:AAC06922.1; PID:g2983333; GB:AE00065  
 A:Experimental source: strain VF5

C:Genetics:

A:Gene: aq\_754

Query Match 0.5%; Score 7; DB 2; Length 99;

Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1260 ELKLAKE 1266

Db 50 ELKLAKE 56

RESULT 31

S72589

hypothetical protein B1937\_F1\_22 - Mycobacterium leprae

C:Species: Mycobacterium leprae

C>Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 22-Oct-1999

C:Accession: S72589

R:Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, November 1993

A:Reference number: S72580

A:Accession: S72589

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-103 <SMI>

A:Cross-references: EMBL:U00016; NID:g466961; PIDN:AA17157.1; PID:g466971

Query Match 0.5%; Score 7; DB 2; Length 103;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 889 IESVFEL 895

Db 2 IESVFEL 8

RESULT 32

B2AG55

virB2 protein precursor - Agrobacterium tumefaciens plasmid pTi15955

C:Species: Agrobacterium tumefaciens

C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999

C:Accession: S00778

R:Thompson, D.V.; Melchers, L.S.; Idler, K.B.; Schilperoort, R.A.; Hooykaas, P.J.J.

Nucleic Acids Res. 16, 4621-4636, 1988

A:Title: Analysis of the complete nucleotide sequence of the Agrobacterium tumefaciens

A:Reference number: S00777; MUID:88247765

A:Accession: S00778

A:Molecule type: DNA

A:Residues: 1-121 <THO>

A:Cross-references: EMBL:X06826; NID:g39195; PIDN:CAA29973.1; PID:g39199.

C:Genetics:

A:Genome: plasmid

C:Superfamily: tumor-inducing plasmid pTiC58 virB2 protein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-121/Product: virB2 protein #status predicted <MAT>

Query Match 0.5%; Score 7; DB 1; Length 121;

Best Local Similarity 100.0%; Pred. No. 59;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 LSLSNAM 975

Db 14 LSLSNAM 20

RESULT 33

B2AG6

virB2 protein - Agrobacterium tumefaciens plasmid pTiA6

C:Species: Agrobacterium tumefaciens

C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999

C:Accession: B28621; B27127

R:Ward, J.E.; Akiyoshi, D.E.; Regier, D.; Datta, A.; Gordon, M.P.; Nester, E.W.

J. Biol. Chem. 263, 5804-5814, 1988

A:Title: Characterization of the virB operon from an Agrobacterium tumefaciens Ti pla

A:Reference number: A28621; MUID:88186901

A:Accession: B28621  
 A:Molecule type: DNA  
 A:Residues: 1-121 <WAR>  
 A:Cross-references: GB:J03216; NID:g1196971; PIDN:AAA88646.1; PID:g1196973  
 C:Genetics:  
 A:Genome: plasmid  
 C:Superfamily: tumor-inducing plasmid pTIC58 virB2 protein

Query Match 0.5%; Score 7; DB 1; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 LSLSNAM 975  
 DB 14 LSLSNAM 20  
 |||||

## RESULT 34

E75555  
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C:Accession: E75555  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896  
 A:Accession: E75555  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-127 <WHI>  
 A:Cross-references: GB:AE001876; GB:AE000513; NID:96457800; PIDN:AAF09726.1; PID:g645780  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR0136  
 A:Map position: 1

Query Match 0.5%; Score 7; DB 2; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 686 LVROGKV 692  
 DB 106 LVROGKV 112  
 |||||

## RESULT 35

T41487  
 very hypothetical protein SPCC622.07 - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T41487  
 R:Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, October 1998  
 A:Reference number: Z21998

A:Accession: T41487  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-128 <SEE>  
 A:Cross-references: EMBL:AL033127; PIDN:CAA21863.1; GSPDB:GN00068; SPDB:SPCC622.07  
 A:Experimental source: strain 972h-; cosmid c622  
 C:Genetics:  
 A:Gene: SPDB:SPCC622.07  
 A:Map position: 3

Query Match 0.5%; Score 7; DB 2; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1267 VFLNLGV 1273  
 DB 65 VFLNLGV 71  
 |||||

## RESULT 36

C83475  
 hypothetical protein PA1354 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: C83475  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;  
 Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic  
 A:Reference number: AB2950; MUID:20437337  
 A:Accession: C83475  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-141 <STO>  
 A:Cross-references: GB:AE004565; GB:AE004091; NID:g9947294; PIDN:AAG04743.1; GSPDB:  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA1354

Query Match 0.5%; Score 7; DB 2; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 VIDGPFA 391  
 DB 64 VIDGPFA 70  
 |||||

## RESULT 37

PQ0506  
 hypothetical protein 151 - fowlpox virus (fragment)  
 N:Alternate names: ORF3 protein  
 C:Species: fowlpox virus  
 C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 03-Nov-2000  
 C:Accession: PQ0506; S27935  
 R:Ogawa, R.; Calvert, J.G.; Yanagida, N.; Nazerian, K.  
 J. Gen. Virol. 74, 55-64, 1993  
 A:Title: Insertional inactivation of a fowlpox virus homologue of the vaccinia viru-  
 A:Reference number: JQ1894; MUID:93139784  
 A:Accession: PQ0506  
 A:Molecule type: DNA  
 A:Residues: 1-151 <OGA>  
 A:Cross-references: GB:M88588; NID:g333522; PIDN:AAA47188.1; PID:g333525  
 A:Note: submitted to the EMBL Data Library, May 1992

Query Match 0.5%; Score 7; DB 2; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 531 GVTOKVN 537  
 DB 69 GVTOKVN 75  
 |||||

## RESULT 38

S33585  
 lipoprotein smpA precursor - Treponema hyodysenteriae  
 C:Species: Treponema hyodysenteriae  
 C:Date: 08-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
 C:Accession: S33585  
 R:Thomas, W.; Sellwood, R.  
 Infect. Immun. 61, 1136-1140, 1993  
 A:Title: Molecular cloning, expression, and DNA sequence analysis of the gene that

A:Reference number: S33585; MUID:93162807

A:Accession: S33585

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-159 <THO>

A:Cross-references: EMBL:X68401; NID:g312941; PIDN:CAA48467.1; PID:g312942

A:Note: The source is designated as Serpulina hyodysenteriae

C:Genetics:

A:Gene: smPA

C:Superfamily: Treponema hyodysenteriae lipoprotein smpA

Query Match 0.5%; Score 7; DB 2; Length 159;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 515 KGIDTCN 521

Db 100 KGIDTCN 106

RESULT 39

JQ1693

pathogenesis-related protein 1 precursor, 17.6K - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Feb-2001

C:Accession: JQ1693; PQ0471; C84519

R:Uknes, S.; Mauch-Mani, B.; Moyer, M.; Potter, S.; Williams, S.; Dincher, S.; Chandler,

Plant Cell 4, 645-656, 1992

A:Title: Acquired resistance in Arabidopsis.

A:Reference number: JQ1693; MUID:93005717

A:Accession: JQ1693

A:Molecule type: mRNA

A:Residues: 1-161 <UKN>

A:Cross-references: GB:M90508; NID:g166860; PIDN:AAA32863.1; PID:g166861

A:Accession: PQ0471

A:Molecule type: protein

A:Residues: 42-67;104-112;153-161 <UKN1>

A:Experimental source: leaf

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: C84519

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <STO>

A:Cross-references: GB:AE002093; NID:g3810599; PIDN:AAC69381.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g14610

A:Map position: 2

C:Superfamily: pathogenesis-related leaf protein

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-161/Product: pathogenesis-related protein 1, 17.6K #status experimental <MAT>

Query Match 0.5%; Score 7; DB 2; Length 161;

Best Local Similarity 100.0%; Pred. No. 77;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGALV 24

Db 15 ALVGALV 21

RESULT 40

S65777

pathogenesis-related protein 1a homolog precursor - rape

C:Species: Brassica napus (rape)

C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Aug-1999

C:Accession: S65777

R:Hanfrey, C.; Fife, M.; Buchanan-Wollaston, V.

Plant Mol. Biol. 30, 597-609, 1996

A:Title: Leaf senescence in Brassica napus: expression of genes encoding pathogenesis

A:Reference number: S65777; MUID:96189271

A:Accession: S65777

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-161 <HAN>

A:Cross-references: EMBL:U21849; NID:g722273; PIDN:AA01666.1; PID:g722274

A:Note: the sequence of residues 1-26 and the corresponding nucleotide sequence are n

C:Superfamily: pathogenesis-related leaf protein

Query Match 0.5%; Score 7; DB 2; Length 161;

Best Local Similarity 100.0%; Pred. No. 77;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGALV 24

Db 15 ALVGALV 21

RESULT 41

H84518

pathogenesis-related PR-1-like protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: H84518

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: H84518

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <STO>

A:Cross-references: GB:AE002093; NID:g3810602; PIDN:AAC69384.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g14580

A:Map position: 2

C:Superfamily: pathogenesis-related leaf protein

Query Match 0.5%; Score 7; DB 2; Length 161;

Best Local Similarity 100.0%; Pred. No. 77;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGALV 24

Db 15 ALVGALV 21

RESULT 42

T45148

methyl coenzyme M reductase (EC 1.8.-.-) II D protein [imported] - Methanobacterium t

C:Species: Methanobacterium thermoautotrophicum

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000

C:Accession: T45148

R:Pihi, T.D.; Sharma, S.; Reeve, J.N.

J. Bacteriol. 176, 6384-6391, 1994

A:Title: Growth phase-dependent transcription of the genes that encode the two methyl

rium thermoautotrophicum delta H

A:Reference number: Z22931; MUID:95014084

A:Accession: T45148

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-161 <PIH>

A:Cross-references: EMBL:U09990; NID:g517420; PIDN:AAA73437.1; PID:g517422

A:Experimental source: delta H

C:Genetics:

A:Gene: mrtD

C:Superfamily: methyl coenzyme M reductase D  
C:Keywords: oxidoreductase

Query Match 0.5%; Score 7; DB 2; Length 161;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 394 KDTVVNI 400  
|||||||  
DB 154 KDTVVNI 160

## RESULT 43

T08154  
pathogenesis-related protein PR1 - rape  
C:Species: Brassica napus (rape)  
C:Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 08-Oct-1999  
C:Accession: T08154  
R:Zhang, P.; Fristensky, B.W.  
submitted to the EMBL Data Library, July 1996  
A:Description: Expression of the Brassica napus PR1 gene is induced during the incompatibility reaction between Brassica napus and the pathogen.  
A:Reference number: 216384  
A:Accession: T08154  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-162 <ZHA>  
A:Cross-references: EMBL:U64806; NID:g1498731; PID:g1498731  
A:Experimental source: cv. Glacier  
C:Genetics:  
A:Gene: Ypr1  
C:Superfamily: pathogenesis-related leaf protein

Query Match 0.5%; Score 7; DB 2; Length 162;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGALV 24  
|||||||  
DB 15 ALVGALV 21

## RESULT 44

F69017  
methyl coenzyme M reductase II, D protein - Methanobacterium thermoautotrophicum (strain F69017)  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1999  
C:Accession: F69017  
R:Smith, D.K.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; K., S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: function and organization of the genome.  
A:Reference number: A69000; MUID:98037514  
A:Accession: F69017  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-162 <MTH>  
A:Cross-references: GB:AE000883; GB:AE000666; NID:g2622231; PIDN:AA85620.1; PID:g2622231  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH131  
C:Superfamily: methyl coenzyme M reductase D

Query Match 0.5%; Score 7; DB 2; Length 162;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 394 KDTVVNI 400  
|||||||  
DB 155 KDTVVNI 161

## RESULT 45

T31280  
benzoate 1,2-dioxygenase (EC 1.14.12.10) homolog - Sphingomonas aromaticivorans plasmid pSph1280  
C:Species: Sphingomonas aromaticivorans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 18-Feb-2000  
C:Accession: T31280  
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.A.  
submitted to the EMBL Data Library, July 1998  
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromaticivorans.  
A:Reference number: Z20992  
A:Accession: T31280  
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-164 <ROM>  
A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378421; PIDN:AA04004.1  
C:Genetics:  
A:Note: xylY  
A:Superfamily: benB protein  
C:Keywords: Oxidoreductase

Query Match 0.5%; Score 7; DB 2; Length 164;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GKVATLS 696  
|||||||  
DB 121 GKVATLS 127

## RESULT 46

D83111  
transcription antitermination protein NusG PA4275 [Imported] - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83111  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.  
A:Reference number: A82950; MUID:20437337  
A:Accession: D83111  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-177 <STO>  
A:Cross-references: GB:AE004843; GB:AE004091; NID:g9950489; PIDN:AAG07663.1; GSPDB:AE004843  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: nusG; PA4275  
C:Superfamily: transcription antitermination factor nusG

Query Match 0.5%; Score 7; DB 2; Length 177;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 VIDGPFA 391  
|||||||  
DB 132 VIDGPFA 138

## RESULT 47

S39859  
transcription antitermination factor nusG - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Jun-2000  
C:Accession: S39859; C69668; S40071  
R:Jeong, S.M.; Yoshikawa, H.; Takahashi, H.  
Mol. Microbiol. 10, 133-142, 1993

A:Title: Isolation and characterization of the secE homologue gene of *Bacillus subtilis*.  
A:Reference number: S39856; MUID:95058172  
A:Accession: S39859  
A:Molecule type: DNA  
A:Residues: 1-177 <JEO>  
A:CROSS-references: ENBL:DL3303; NID:9285624; PIDN:BA02560.1; PID:9285628  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galler  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A:Reference number: A69580; MUID:98044033  
A:Accession: C69688  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-177 <KUN>  
A:CROSS-references: GB:299104; GB:AL009126; NID:92632267; PIDN:CAB11877.1; PID:92632368  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: nusG  
C:Superfamily: transcription antitermination factor nusG  
C:Keywords: transcription antitermination

Query Match 0.5%; Score 7; DB 2; Length 177;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 VIDGPFA 391  
|||||  
Db 132 VIDGPFA 138

RESULT 48  
E64678  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain N0010 - *Helicobacter pylori* (strain 2  
C:Species: *Helicobacter pylori*  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: E64678  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A:Reference number: A64520; MUID:97394467  
A:Accession: E64678  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-182 <TOM>  
A:CROSS-references: GB:AE000631; GB:AE000511; NID:92314421; PIDN:AAD08313.1; PID:9231443  
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 6  
C:Keywords: electron transfer; membrane-associated complex; NAD; oxidoreductase

Query Match 0.5%; Score 7; DB 2; Length 182;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 VGGIATG 51  
|||||  
Db 159 VGGIATG 165

RESULT 49  
C71839  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain J - *Helicobacter pylori* (strain.  
C:Species: *Helicobacter pylori*  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
C:Accession: C71839  
R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.  
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MUID:99120557  
A:Accession: C71839  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-182 <ARN>  
A:CROSS-references: GB:AE001545; GB:AE001439; NID:94155776; PIDN:AAD06756.1; PID:9415  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: nuoJ  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 6  
C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 0.5%; Score 7; DB 2; Length 182;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 VGGIATG 51  
|||||  
Db 159 VGGIATG 165

RESULT 50  
F82285  
conserved hypothetical protein VC0747 [imported] - *Vibrio cholerae* (strain N16961 ser  
C:Species: *Vibrio cholerae*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: F82285  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers  
L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: A82035; MUID:20406833  
A:Accession: F82285  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-188 <HEI>  
A:CROSS-references: GB:AE004160; GB:AE003852; NID:9655181; PIDN:AAF93912.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0747  
A:Map position: 1  
C:Superfamily: hypothetical protein b2531

Query Match 0.5%; Score 7; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 360 SQNNSNT 366  
|||||  
Db 168 SQNNSNT 174

RESULT 51  
S08632  
nodulin-21 - soybean  
C:Species: Glycine max (soybean)  
C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 15-Sep-2000  
C:Accession: S08632

R: Delauney, A.J.; Cheon, C.I.; Snyder, P.J.; Verma, D.P.S.  
Plant Mol. Biol. 14, 449-451, 1990

A: Title: A nodule-specific sequence encoding a methionine-rich polypeptide, nodulin-21.

A: Reference number: S08632; MUID: 91346633

A: Accession: S08632

A: Molecule type: mRNA

A: Residues: 1-206 <DEL>

A: Cross-references: EMBL: X16488; NID: g18693; PIDN: CAA34506.1; PID: g18694  
C: Superfamily: Streptomyces coelicolor probable membrane protein SC3A3.05c

Query Match 0.5%; Score 7; DB 2; Length 206;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 789 ACGMAIG 795

DB 100 ACGMAIG 106

RESULT 52

D84601

phosphoserine phosphatase (EC 3.1.3.3) - Helicobacter pylori (strain 26695)

C: Species: Helicobacter pylori

C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C: Accession: D84601

R: Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A: Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A: Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A: Reference number: A64520; MUID: 97394467

A: Accession: D84601

A: Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-207 <TON>

A: Cross-references: GB: AE000578; GB: AE000511; NID: g2313759; PIDN: AAD07711.1; PID: g231377

C: Genetics:

A: Start codon: GTG

C: Superfamily: phosphoserine phosphatase

C: Keywords: phosphoric monoester hydrolase

Query Match 0.5%; Score 7; DB 1; Length 207;

Best Local Similarity 100.0%; Pred. No. 97;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1261 LKLAKEV 1267

DB 62 LKLAKEV 68

RESULT 53

C71914

phosphoserine phosphatase - Helicobacter pylori (strain J99)

C: Species: Helicobacter pylori

A: Variety: strain J99

C: Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 22-Jun-1999

C: Accession: C71914

R: Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jlang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A: Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A: Reference number: A71800; MUID: 99120557

A: Accession: C71914

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-207 <ARN>

A: Cross-references: GB: AE001491; GB: AE001439; NID: g4155127; PIDN: AAD06170.1; PID: g415514

A: Experimental source: strain J99

C: Genetics:

A: Gene: serB

C: Superfamily: phosphoserine phosphatase

Query Match 0.5%; Score 7; DB 2; Length 207;

Best Local Similarity 100.0%; Pred. No. 97;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1261 LKLAKEV 1267

DB 62 LKLAKEV 68

RESULT 54

S78275

ribosomal protein L4, chloroplast - Odontella sinensis chloroplast

C: Species: chloroplast Odontella sinensis

C: Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 20-Jun-2000

C: Accession: S78275

R: Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A: Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella si

A: Reference number: S78238

A: Accession: S78275

A: Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-215 <KOW>

A: Cross-references: EMBL: 267753; NID: g1185127; PIDN: CAA91648.1; PID: g1185165

A: Note: the nucleotide sequence was submitted to the EMBL Data Library, November 19;

C: Genetics:

A: Gene: rpl4

A: Genome: chloroplast

C: Superfamily: Escherichia coli ribosomal protein L4

C: Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 0.5%; Score 7; DB 2; Length 215;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 624 RNIRNVE 630

DB 175 RNIRNVE 181

RESULT 55

F75166

hypothetical protein PAB0330 - Pyrococcus abyssi (strain Orsay)

C: Species: Pyrococcus abyssi

C: Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C: Accession: F75166

R: anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A: Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome

A: Reference number: A75001

A: Accession: F75166

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-215 <RAW>

A: Cross-references: GB: AJ248284; GB: AL096836; NID: g5457730; PIDN: CAB49413.1; PID: g54

A: Experimental source: strain Orsay

C: Genetics:

A: Gene: PAB0330

Query Match 0.5%; Score 7; DB 2; Length 215;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1063 VEALVGG 1069

DB 186 VEALVGG 192

## RESULT 56

T17795

Hypothetical protein A298L - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T17795

R:Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A:Reference number: 218806

A:Accession: T17795

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-225 &lt;GRA&gt;

A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AA096666.1

A:Experimental source: specific host Chlorella strain NC64A

C:Genetics:

A:Note: A298L

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 225;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 600 VKFKGGE 606

Db 21 VKFKGGE 27

## RESULT 57

A53273

MHC class II histocompatibility antigen DR alpha chain - horse (fragment)

C:Species: Equus caballus (domestic horse)

C:Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 21-Jan-2000

C:Accession: A53273

R:Albright, D.; Bailey, E.; Woodward, J.G.

Immunogenetics 34, 136-138, 1991

A:Title: Nucleotide sequence of a cDNA clone of the horse (Equus caballus) DR alpha gene.

A:Reference number: A53273; MUID:91331619

A:Accession: A53273

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-226 &lt;ALB&gt;

A:Cross-references: GB:M60100; NID:g164236; PIDN:AAA30956.1; PID:g164237

C:Genetics:

A:Gene: DRA

A:Map position: 20

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

F:97-162/Domain: immunoglobulin homology &lt;IMM&gt;

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 226;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 FEAQAL 1115

Db 51 FEAQAL 57

## RESULT 58

I54426

MHC H2-w28-E alpha - ricefield mouse (Mus caroli) (fragment)

C:Species: Mus caroli

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jan-2000

C:Accession: I54426

R:Kasahara, M.; Stojiljkovic, I.; Mayer, W.E.; Dembic, Z.; Figueroa, F.; Klein, J.

Immunogenetics 24, 324-327, 1996

A:Title: The nucleotide sequence of the mouse H-2Ew28-alpha gene.

A:Reference number: I54426; MUID:87055942

A:Accession: I54426

A&gt;Status: preliminary;

A:Molecule type: DNA

A:Residues: 1-227 &lt;RES&gt;

A:Cross-references: GB:M16241; NID:g199685; PIDN:AAA39705.1; PID:g199688

C:Genetics:

A:Introns: 82/1; 176/1

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

F:97-162/Domain: immunoglobulin homology &lt;IMM&gt;

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 227;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 FEAQAL 1115

Db 51 FEAQAL 57

## RESULT 59

G69491

conserved hypothetical protein AF1936 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000

C:Accession: G69491

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod

Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A:Reference number: A69250; MUID:98049343

A:Accession: G69491

A&gt;Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-227 &lt;KLE&gt;

A:Cross-references: GB:AE000970; GB:AE000782; NID:g2689293; PIDN:AAB89319.1; PID:g264

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1839

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 227;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1011 ESAAEVL 1017

Db 28 ESAAEVL 34

## RESULT 60

A48381

MHC class II histocompatibility antigen alpha chain - sheep

C:Species: Ovis sp. (sheep)

C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000

C:Accession: A48381

R:Ballingall, K.T.; Wright, H.; Redmond, J.; Dutia, B.M.; Hopkins, J.; Lang, J.; Deve

Anim. Genet. 23, 347-359, 1992

A:Title: Expression and characterization of ovine major histocompatibility complex cl

A:Reference number: A48381; MUID:92367958

A:Accession: A48381

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229 &lt;BAL&gt;

A:Experimental source: Suffolk

A:Note: sequence extracted from NCBI backbone (NCBIN:110906, NCBIN:110908, NCBI:1109

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

F:100-165/Domain: immunoglobulin homology &lt;IMM&gt;

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 229;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 FEAQAL 1115

Db 51 FEAQAL 57

Db 54 FEAQAL 60

## RESULT 61

MHC H2-IE-alpha cell surface glycoprotein - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 21-Jan-2000  
 C:Accession: I55971  
 R:Ayane, M.; Mengle-Gaw, L.; McDevitt, H.O.; Benoist, C.O.; Mathis, D.  
 J. Immunol. 137, 948-951, 1986  
 A:Title: E-alpha-u and E-beta-u chain association: Where lies the anomaly?  
 A:Reference number: I55971; MUID:86252293  
 A:Accession: I55971  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-229 <RES>  
 A:Cross-references: GB:M12818; NID:g199500; PIDN:AAA39638.1; PID:g199501  
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
 C:Keywords: glycoprotein  
 F:99-164/Domain: Immunoglobulin homology <IMM>

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 229;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115

Db 53 FEAQAL 59

## RESULT 62

T27840  
 hypothetical protein ZK39.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T27840  
 R:Kershaw, J.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z20428  
 A:Accession: T27840  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-229 <WIL>  
 A:Cross-references: EMBL:Z82093; PIDN:CAB05018.1; GSPDB:GN00019; CESP:ZK39.2  
 A:Experimental source: clone ZK39  
 C:Genetics:  
 A:Gene: CESP:ZK39.2  
 A:Map position: 1  
 A:Introns: 28/1; 55/3; 111/1

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 229;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 TLGLRN 150

Db 96 TLGLRN 102

## RESULT 63

H69488  
 SSU ribosomal protein S4E (rps4E) homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 13-Aug-1999  
 C:Accession: H69488  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing ar.  
 A:Reference number: A69250; MUID:98049343  
 A:Accession: H69488  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-235 <KLE>  
 A:Cross-references: GB:A6000971; GB:AE000782; NID:g2689294; PIDN:AAB89340.1; PID:g2.  
 C:Superfamily: rat ribosomal protein S4

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 235;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 686 LVRGKV 692

Db 125 LVRGKV 131

## RESULT 64

I67432

BCL-2 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 16-Jul-1999

C:Accession: I67432

R:Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.

Endocrinology 136, 232-241, 1995

A:Title: Expression of members of the bcl-2 gene family in the immature rat ovary:

constitutive bcl-2 and bcl-x-long messenger ribonucleic acid levels.

A:Reference number: I53295; MUID:95129487.

A:Accession: I67432

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-236 &lt;RES&gt;

A:Cross-references: EMBL:U34964; NID:g1004378; PIDN:AAA77687.1; PID:g1004379

C:Superfamily: bcl transforming protein

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 236;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVA 22

Db 219 SLALVA 225

## RESULT 65

I53744

gene bcl-2 protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 16-Jul-1999

C:Accession: I53744

R:Sato, T.; Irie, S.; Krajewski, S.; Reed, J.C.

Gene 140, 291-292, 1994

A:Title: Cloning and sequencing of a cDNA encoding the rat Bcl-2 protein.

A:Reference number: I53744; MUID:94193015

A:Accession: I53744

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-236 &lt;RES&gt;

A:Cross-references: GB:L14680; NID:g408946; PIDN:AAA53662.1; PID:g408947

C:Genetics:

A:Gene: bcl-2

C:Superfamily: bcl transforming protein

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 236;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVA 22



```

Db 219 SLALVGA 225
|||||||
RESULT 66
JC7383
B-cell lymphoma 2 protein - Chinese hamster
C:Species: Cricetus griseus (Chinese hamster)
C>Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 08-Dec-2000
C:Accession: JC7383
R:Tomicic, M.T.; Christmann, M.; Kaina, B.
Biochem. Biophys. Res. Commun. 275, 899-903, 2000
A:Title: Cloning and functional analysis of cDNA encoding the hamster Bcl-2 protein.
A:Reference number: JC7383
A:Contents: Ovary
A:Accession: JC7383
A:Molecule type: mRNA
A:Residues: 1-236 <TOM>
A:Cross-references: GB:AJ271720
C:Comment: This protein has anti-apoptotic function, and supports cell survival.
C:Genetics:
A:Gene: bcl-2
C:Superfamily: bcl transforming protein
C:Keywords: B-cell lymphoma; ovary

Query Match 0.5%; Score 7; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22
|||||||
Db 219 SLALVGA 225

RESULT 67
TVHUAL
Transforming protein bcl-2, splice form alpha - human
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1988 #sequence_revision 07-Jun-1996 #text_change 15-Oct-1999
C:Accession: C37332; A29409; S02452; A24428; A27622; B27622
R:Eguchi, Y.; Evert, D.L.; Tsujimoto, Y.
Nucleic Acids Res. 20, 4187-4192, 1992
A:Title: Isolation and characterization of the chicken bcl-2 gene: expression in a variety of tissues.
A:Reference number: A37332; MUID:92375724
A:Accession: C37332
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-239 <EGU>
A:Note: this report is a correction
R:Tsujimoto, Y.; Croce, C.M.
Proc. Natl. Acad. Sci. U.S.A. 83, 5214-5218, 1986
A:Title: Analysis of the structure, transcripts, and protein products of bcl-2, the gene for the B-cell lymphoma.
A:Reference number: A29409; MUID:86259760
A:Accession: A29409
A:Molecule type: mRNA
A:Residues: 1-95, 'A', 97-109, 'G', 111-236, 'S', 238-239 <TSD>
A:Cross-references: GB:M13994; MUID:9179366; PIDN:AAA51813.1; PID:9179367
A:Note: This sequence has been corrected in reference A37332
R:Seto, M.; Jaeger, U.; Hockett, R.D.; Graninger, W.; Bennett, S.; Goldman, P.; Korsmeyer, S.J.
EMBO J. 7, 123-131, 1988
A:Title: Alternative promoters and exons, somatic mutation and deregulation of the Bcl-2 gene.
A:Reference number: S02452; MUID:88196071
A:Accession: S02452
A:Molecule type: mRNA
A:Residues: 1-239 <SET>
R:Cleary, M.L.; Smith, S.D.; Sklar, J.
Cell 47, 19-28, 1986
A:Title: Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin heavy chain enhancer of the 3' long terminal repeat.
A:Reference number: A24428; MUID:87002488
A:Accession: A24428
A:Molecule type: mRNA
A:Residues: 1-58, 'T', 60-116, 'R', 118-239 <CLE>

A:Cross-references: GB:M14745; MUID:9179370; PIDN:AAA35591.1; PID:9179371
R:Hua, C.; Zorn, S.; Jensen, J.P.; Coupland, R.W.; Ko, H.S.; Wright, J.J.; Bakhshi, A.
Oncogene Res. 2, 263-275, 1988
A:Title: Consequences of the t(14;18) chromosomal translocation in follicular lymphoma
A:Reference number: A27622; MUID:88217344
A:Accession: A27622
A:Molecule type: mRNA
A:Residues: 1-58, 'T', 60-239 <HUA>
A:Accession: B27622
A:Molecule type: DNA
A:Residues: 1-6, 'S', 8-58, 'T', 60-128, 'C', 130-239 <HUA2>
A:Note: the sequence was determined from the germline gene
C:Comment: Constitutive expression of BCL2 following t(14;18) chromosomal translocation
C:Genetics:
A:Gene: GDB:BCL2
A:Cross-references: GDB:119031; OMIM:151430
A:Map position: 18q21.3-18q21.3
C:Function:
A:Description: blocks apoptosis in hematopoietic cells
C:Superfamily: bcl transforming protein
C:Keywords: alternative splicing; apoptosis; B-cell lymphoma; follicular lymphoma; pr

Query Match 0.5%; Score 7; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22
|||||||
Db 222 SLALVGA 228

RESULT 68
D72747
probable proteasome, beta subunit APE0507 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: D72747
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: D72747
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-239 <KAW>
A:Cross-references: DDBJ:AF000059; MUID:95103911; PIDN:BAA79472.1; PID:95104156
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0507
C:Superfamily: multicatalytic endopeptidase complex chain C9

Query Match 0.5%; Score 7; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1063 VEAIVGG 1069
|||||||
Db 138 VEAIVGG 144

RESULT 69
S37342
chitinase (EC 3.2.1.14) ch17 precursor - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S37342; S38837; S25635
R:Danhush, N.; Wagemakers, C.A.M.; van Kan, J.A.L.; de Wit, P.J.G.M.
Plant Mol. Biol. 22, 1017-1029, 1993
A:Title: Molecular characterization of four chitinase cDNAs obtained from Cladosporiu
A:Reference number: S37341; MUID:94003061

```

A:Accession: S37342  
A:Molecule type: mRNA  
A:Residues: 1-247 <DAN>  
A:Cross-references: EMBL:Z15139; NID:gl9186; PIDN:CAA78844.1; PID:g19187  
A:Accession: S38837

A:Molecule type: protein  
A:Residues: 112-115;183-188;218-221 <DA2>

C:Genetics:

A:Gene: ch17

C:Superfamily: plant chitinase; plant chitinase homology  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-247/Product: chitinase ch17 #status predicted <MAT>  
F:22-245/Domain: plant chitinase homology <PCH>

Query Match 0.5%; Score 7; DB 2; Length 247;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 DGPFAGG 393

Db 93 DGPFAGG 99

RESULT 70

A71684

probable conjugal transfer protein trbg precursor (trbg) RP286 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000

C:Accession: A71684

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499

A:Accession: A71684

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-250 <AND>

A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14747.1; PID:g386084

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: trbg; RP286

Query Match 0.5%; Score 7; DB 2; Length 250;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 DKNABIS 237

Db 190 DKNABIS 196

RESULT 71

A46505

SFA-DRAD (MHC Class II) - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 18-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000

C:Accession: A46505

R:Hirsch, F.; Germana, S.; Gustafsson, K.; Pratt, K.; Sachs, D.H.; Leguern, C.

J. Immunol. 149, 841-846, 1992

A:Title: Structure and expression of class II alpha genes in miniature swine.

A:Reference number: A46505; MUID:92340887

A:Accession: A46505

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-252 <HIR>

A:Cross-references: GB:M93028; NID:g164551; PIDN:AAA31075.1; PID:g164552

A:Note: sequence extracted from NCBI backbone (NCBIN:109901, NCBIPI:109902)

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

F:123-188/Domain: immunoglobulin homology <IMM>

Query Match 0.5%; Score 7; DB 2; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAGAL 1115

Db 77 FEAQAGAL 83

RESULT 72

A83089

conserved hypothetical protein PA4445 [imported] - Pseudomonas aeruginosa (strain P.

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: A83089

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic i

A:Reference number: A82950; MUID:20437337

A:Accession: A83089

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-252 <STO>

A:Cross-references: GB:AE004859; GB:AE004091; NID:g9950678; PIDN:AAG07833.1; GSPDB:

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA4445

C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1777

Query Match 0.5%; Score 7; DB 2; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053-AGVDAYL 1059

Db 190 AGVDAYL 196

RESULT 73

C83778

ferrichrome ABC transporter (ATP-binding protein) BH1027 [imported] - Bacillus halo

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000

C:Accession: C83778

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.;

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans

A:Reference number: A83650; MUID:20263314

A:Accession: C83778

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-252 <STO>

A:Cross-references: GB:AP001510; GB:BA000004; NID:gl0173440; PIDN:BA004746.1; GSPDB

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH1027

Query Match 0.5%; Score 7; DB 2; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 GEKLVID 611

Db 12 GEKLVID 18

RESULT 74

S15684

MHC class II histocompatibility antigen Ovar-DR alpha chain - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jan-2000  
 C:Accession: I47075; S15684  
 R:Fabn, S.A.; Maddox, J.F.; Gogolin-Ewens, K.J.; Baker, L.; Wu, M.J.; Brandon, M.R.  
 Anim. Genet. 24, 249-255, 1993  
 A:Title: Isolation, characterization and evolution of ovine major histocompatibility complex class II genes  
 A:Reference number: I47075; MUID:94057592  
 A:Accession: I47075  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-253 <FA2>  
 A:Cross-references: GB:M73983; NID:g165867; PIDN:AAA16793.1; PID:g165868; EMBL:X60515  
 C:Genetics:  
 A:Gene: MHC Ovar-DR  
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
 F:124-189/Domain: immunoglobulin homology <IMM>

Query Match 0.5%; Score 7; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 |||||  
 Db 78 FEAQAL 84

RESULT 75  
 JC2388  
 class II histocompatibility antigen DR alpha chain (clone W3) precursor - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 28-Feb-1995 #sequence\_revision 03-Apr-1995 #text\_change 20-Jun-2000  
 C:Accession: JC2388; A37206  
 R:Aida, Y.; Kohda, C.; Morooka, A.; Nakai, Y.; Ogimoto, K.; Urao, T.; Asahina, M.  
 Biochem. Biophys. Res. Commun. 204, 195-202, 1994  
 A:Title: Cloning of cDNAs and the molecular evolution of a bovine MHC class II DR gene.  
 A:Reference number: JC2388; MUID:95032095  
 A:Accession: JC2388  
 A:Molecule type: mRNA  
 A:Residues: 1-253 <AID>  
 A:Cross-references: DDBJ:D37956; NID:g790949; PIDN:BA07173.1; PID:g790950  
 A:Experimental source: lymphoid cell line BLSC-KU-1  
 R:van der Poel, J.J.; Groenen, M.A.M.; Dijkhof, R.J.M.; Ruyter, D.; Giphart, M.J.  
 Immunogenetics 31, 29-36, 1990  
 A:Title: The nucleotide sequence of the bovine MHC class II alpha genes: DRA, DQA, and DQ  
 A:Reference number: A37206; MUID:90129153  
 A:Accession: A37206  
 A:Molecule type: DNA  
 A:Residues: 28-253 <VAN>  
 A:Cross-references: GB:M30120; NID:g163370; PIDN:AAA30645.1; PID:g163371  
 C:Genetics:  
 A:Introns: 82/1; 176/1  
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
 C:Keywords: glycoprotein; transmembrane protein  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-253/Product: class II antigen DR alpha chain, major histocompatibility complex #status p  
 F:25-108/Product: alpha 1 #status predicted <AP1>  
 F:109-202/Product: alpha 2 #status predicted <AP2>  
 F:124-189/Domain: immunoglobulin homology <IMM>  
 F:203-215/Domain: connecting #status predicted <CNE>  
 F:216-239/Domain: transmembrane #status predicted <TMW>  
 F:240-253/Domain: intracellular #status predicted <INT>  
 F:102/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 0.5%; Score 7; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 |||||  
 Db 78 FEAQAL 84

## RESULT 76

HLHDA  
 MHC class II histocompatibility antigen HLA-DR alpha chain precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 17-Dec-1982 #sequence\_revision 27-Nov-1985 #text\_change 01-Dec-2000  
 C:Accession: A93922; A20898; A21113; I58045; A91707; A90825; B90825; A93927; I52975;  
 R:Das, H.K.; Lawrence, S.K.; Weissman, S.M.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 3543-3547, 1983  
 A:Title: Structure and nucleotide sequence of the heavy chain gene of HLA-DR.  
 A:Reference number: A93952; MUID:83221632  
 A:Accession: A93952  
 A:Molecule type: DNA  
 A:Residues: 1-254 <DAS>  
 A:Cross-references: GB:J00203; GB:J00204; NID:g188427; PIDN:AAA36302.1; PID:g386945  
 A:Note: this allele is designated DRA\*0101  
 R:Schamboeck, A.; Korman, A.J.; Kamb, A.; Strominger, J.L.  
 Nucleic Acids Res. 11, 8663-8675, 1983  
 A:Title: Organization of the transcriptional unit of a human class II histocompatibility  
 A:Reference number: A20898; MUID:84169507  
 A:Accession: A20898  
 A:Molecule type: DNA  
 A:Residues: 1-241, 'L', 243-254 <SCH>  
 A:Experimental source: (unknown allele type)  
 R:Das, H.K.; Biro, P.A.; Cohen, S.N.; Erlich, H.A.; von Gabain, A.; Lawrence, S.K.; L  
 Proc. Natl. Acad. Sci. U.S.A. 80, 1531-1535, 1983  
 A:Title: Use of synthetic oligonucleotide probes complementary to genes for human HLA  
 A:Reference number: A21113; MUID:83169718  
 A:Accession: A21113  
 A:Molecule type: mRNA  
 A:Residues: 1-39 <DAS>  
 A:Cross-references: GB:J00197  
 R:Lee, J.S.; Trowsdale, J.; Travers, P.J.; Carey, J.; Grosveld, F.; Jenkins, J.; Bodm  
 Nature 299, 750-752, 1982  
 A:Title: sequence of an hla-dr alpha-chain cdna clone and intron-exon organization of  
 A:Reference number: I58045; MUID:83013020  
 A:Accession: I58045  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-254 <RES>  
 A:Cross-references: GB:J00194; NID:g188231; PIDN:AAA36275.1; PID:g307264  
 R:Das, H.K.; Lawrence, S.K.; Weissman, S.M.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 7024, 1983  
 A:Reference number: A93978  
 A:Contents: annotation; erratum  
 R:Yang, C.Y.; Kratzin, H.; Gotz, H.; Thinnies, F.P.; Kruse, T.; Egert, G.; Pauly, E.;  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 671-676, 1982  
 A:Title: Primärstruktur menschlicher Histokompatibilitätsantigene der Klasse II. 2.  
 A:Reference number: A91707; MUID:82263347  
 A:Accession: A91707  
 A:Molecule type: protein  
 A:Residues: 26-148, 'D', 150-204 <YAN>  
 R:Larhammar, D.; Gustafsson, K.; Claesson, L.; Bill, P.; Wiman, K.; Schenning, L.; Su  
 Cell 30, 153-161, 1982  
 A:Title: Alpha chain of HLA-DR transplantation antigens is a member of the same prote  
 A:Reference number: A90825; MUID:83025073  
 A:Accession: A90825  
 A:Molecule type: protein  
 A:Residues: 26-60 <LAR>  
 A:Note: 28-Ala, 29-Asp, 33-Thr, 33-Pro, 34-Tyr, 35-Pro, 48-Gln, and 54-Thr were also  
 A:Accession: B90825  
 A:Molecule type: mRNA  
 A:Residues: 32-202, 204-254 <LA2>  
 A:Cross-references: GB:J00196  
 R:Korman, A.J.; Auffray, C.; Schamboeck, A.; Strominger, J.L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 6013-6017, 1982  
 A:Title: The amino acid sequence and gene organization of the heavy chain of the HLA-  
 A:Reference number: A93927; MUID:83299916  
 A:Accession: A93927  
 A:Molecule type: DNA  
 A:Residues: 29-254 <KOR>

A:Cross-references: GB:J00201  
A:Note: this allele was also found  
R:Kajimura, Y.; Toyoda, H.; Sato, M.; Miyakoshi, S.; Kaplan, S.A.; Ike, Y.; Goyert, S.M.  
DNA 2, 175-182, 1983  
A:Title: Cloning the heavy chain of human HLA-DR antigen using synthetic oligodeoxyribonucleotide  
A:Accession: 152975  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-254 <KAY>  
A:Cross-references: GB:K01171; NID:g188264; PIDN:AAA59785.1; PID:g307267  
R:Gustafsson, K.; Wiman, K.; Larhammar, D.G.; Rask, L.; Peterson, P.A.  
Scand. J. Immunol. 19, 91-97, 1984  
A:Title: Signal sequences distinguish class II histocompatibility antigen beta chains of HLA-DR and HLA-DQ  
A:Reference number: 159467; MUID:84146572  
A:Accession: 180355  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-50 <RE2>  
A:Cross-references: GB:M35979; NID:g188262; PIDN:AAA36283.1; PID:g188263  
R:Lee, J.S.; Irowdale, J.; Bodmer, W.F.  
Proc. Natl. Acad. Sci. U.S.A. 79, 545-549, 1982  
A:Title: cDNA clones coding for the heavy chain of human hla-dr antigen.  
A:Reference number: 158984; MUID:82197531  
A:Accession: 158984  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 26-42 <RE3>  
A:Cross-references: GB:J00193; NID:g188213; PIDN:AAA36272.1; PID:g188214  
R:Koppelman, B.; Cresswell, P.  
J. Immunol. 145, 2730-2736, 1990  
A:Title: Rapid nonlysosomal degradation of assembled HLA class II glycoproteins incorporated into lipid vesicles  
A:Reference number: 156085; MUID:91010755  
A:Accession: 156085  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-241, 'L', 243-254 <RE4>  
A:Cross-references: GB:M60334; NID:g188255; PIDN:AAA59783.1; PID:g188256  
R:Korman, A.J.; Knudsen, P.J.; Kaufman, J.F.; Strominger, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 1844-1848, 1982  
A:Title: cDNA clones for the heavy chain of HLA-DR antigens obtained after immunopurification from human monocytes  
A:Reference number: 137530; MUID:82197594  
A:Accession: 137530  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 224-241, 'L', 243-254 <RE6>  
A:Cross-references: EMBL:V00528; NID:g32192; PIDN:CAA23787.1; PID:g825675  
C:Genetics:  
A:Gene: GDB:HLA-DRA  
A:Cross-references: GDB:120641; OMIM:142860  
A:Map position: 6p21.3-6p21.3  
A:Introns: 82/1; 176/1  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-254/Product: class II histocompatibility antigen HLA-DR alpha chain #status predicted <TRM>  
F:26-216/Domain: extracellular #status predicted <EXT>  
F:26-109/Domain: alpha-1 <EX1>  
F:125-190/Domain: immunoglobulin homology <IMM>  
F:217-239/Domain: transmembrane #status predicted <TSM>  
F:240-254/Domain: intracellular #status predicted <INT>  
F:103,143/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:132-188/Disulfide bonds: #status experimental

Query Match 0.5%; Score 7; DB 1; Length 254;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAGAL 1115  
|||||  
DB 79 FEAGAL 85

## RESULT 77

HLMSD  
H-2 class II histocompatibility antigen E-d alpha chain precursor - mouse  
N:Alternate names: immune response protein I-E-alpha(d)  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Apr-1984 #sequence\_revision 18-Apr-1984 #text\_change 22-Jun-1999  
C:Accession: B91743; A94266; A93967; A21217; S20788; A02207  
R:Larhammar, D.; Andersson, G.; Andersson, M.; Bill, P.; Bohme, J.; Claesson, L.; De  
Ivenius, B.; Widmark, E.; Rask, L.; Peterson, P.A.  
Hum. Immunol. 8, 95-103, 1983  
A:Title: Molecular analysis of human class II transplantation antigens and their genes  
A:Reference number: A91743; MUID:84031733  
A:Accession: B91743  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-255 <LAR>  
R:McNicholas, J.; Steinmetz, M.; Hunkapiller, T.; Jones, P.; Hood, L.  
Science 218, 1229-1232, 1982  
A:Title: DNA sequence of the gene encoding the E-alpha Ia polypeptide of the BALB/c  
A:Reference number: A94266; MUID:83067428  
A:Accession: A94266  
A:Molecule type: DNA  
A:Residues: 23-154, 'D', 203-238, 'M', 240-255 <MCN>  
R:Benoist, C.O.; Mathis, D.J.; Kanter, M.R.; Williams II, V.E.; McDewitt, H.O.  
Proc. Natl. Acad. Sci. U.S.A. 80, 534-538, 1983  
A:Title: The murine Ia alpha chains, E-alpha and A-alpha, show a surprising degree of  
A:Reference number: A93967; MUID:83169693  
A:Accession: A93967  
A:Molecule type: mRNA; DNA  
A:Residues: 1-201, 'H', 203-218, 'V', 220-238, 'M', 240-255 <BEN>  
R:Hyldig-Nielsen, J.J.; Schenning, L.; Hammerling, U.; Widmark, E.; Heidin, E.; Lind  
Peterson, P.A.; Rask, L.  
Nucleic Acids Res. 11, 5055-5071, 1983  
A:Title: The complete nucleotide sequence of the I-Ealpha(d) immune response gene.  
A:Reference number: A21217; MUID:83272951  
A:Accession: A21217  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-238, 'M', 240-255 <HYL>  
A:Cross-references: GB:K00971; NID:g199498; PIDN:AAA98624.1; PID:g387465  
R:Nygard, N.R.; McCarthy, D.M.; Schiffenauer, J.; Schwartz, B.D.  
Submitted to the EMBL Data Library, August 1990  
A:Description: Nucleotide sequence of MHC class II genes in the NZB mouse.  
A:Reference number: S20786  
A:Accession: S20786  
A:Molecule type: DNA  
A:Residues: 29-109 <NYG>  
A:Cross-references: EMBL:X54427; NID:g53097; PIDN:CAA38299.1; PID:g53098  
C:Genetics:  
A:Introns: 110/1; 204/1  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:125-190/Domain: immunoglobulin homology <IMM>  
F:217-239/Domain: transmembrane #status predicted <TRM>  
F:132-188/Disulfide bonds: #status experimental  
F:143/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 0.5%; Score 7; DB 1; Length 255;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAGAL 1115  
|||||  
DB 79 FEAGAL 85

## RESULT 78

HLMSA  
H-2 class II histocompatibility antigen E-k alpha chain precursor - mouse

C:Species: Mus musculus (house mouse)  
C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 20-Mar-1998  
C:Accession: A21938; A02208  
R:Mathis, D.J.; Benoist, C.O.; Williams II, V.E.; Kanter, M.R.; McDevitt, H.O.  
Cell 32, 745-754, 1983  
A:Title: The murine E-alpha immune response gene.  
A:Reference number: A21938; MUID:83155651  
A:Accession: A21938  
A:Molecule type: DNA  
A:Residues: 1-255 <MATH>  
A:Cross-references: GB:J00398; NID:g199348; PID:g387448  
R:Benoist, C.O.; Mathis, D.J.; Kanter, M.R.; Williams II, V.E.; McDevitt, H.O.  
Proc. Natl. Acad. Sci. U.S.A. 80, 534-538, 1983  
A:Title: The murine Ia alpha chains, E-alpha and A-alpha, show a surprising degree of se  
A:Reference number: A93967; MUID:83169693  
A:Accession: A02208  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-255 <BEN>  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
C:Keywords: heterodimer; transmembrane protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-255/Product: H-2 class II histocompatibility antigen E-k alpha chain #status predic  
F:110-203/Domain: alpha-1 <XAL>  
F:125-190/Domain: immunoglobulin homology <IMM>  
F:204-216/Domain: connecting peptide #status predicted <CCP>  
F:217-244/Domain: transmembrane #status predicted <TM>  
F:245-255/Domain: intracellular #status predicted <INT>  
F:132-188/Disulfide bonds: #status predicted

Query Match 0.5%; Score 7; DB 1; Length 255;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
Db 79 FEAQAL 85

RESULT 79  
S06316  
class II histocompatibility antigen RT1-D alpha(u) chain precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 21-Jan-2000  
C:Accession: S06316  
R:Holowachuk, E.W.; Greer, M.K.; Martin, D.R.  
Nucleic Acids Res. 15, 10551-10567, 1987  
A:Title: The complete sequence of the MHC class II chain RT1D-alpha(u) of the diabetic B  
A:Reference number: S06316; MUID:88096585  
A:Accession: S06316  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-255 <HOL>  
A:Cross-references: GB:Y00480; NID:g57163; PIDN:CAA68540.1; PID:g57164  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-255/Product: class II histocompatibility antigen, RT1-D alpha(u) chain #status pred  
F:110-203/Domain: extracellular alpha-1 #status predicted <ACH1>  
F:125-190/Domain: extracellular alpha-2 #status predicted <ACH2>  
F:204-216/Domain: immunoglobulin homology <IMM>  
F:217-239/Domain: connecting peptide #status predicted <CCP>  
F:240-255/Domain: transmembrane #status predicted <TM>  
F:240-255/Domain: intracellular #status predicted <INT>

Query Match 0.5%; Score 7; DB 2; Length 255;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115

Db 79 FEAQAL 85

RESULT 80  
A45881  
MHC class II histocompatibility antigen RLA-DR alpha chain precursor - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 21-Jan-2000  
C:Accession: A45881  
R:Laverriere, A.; Kulaga, H.; Kindt, T.J.; LeGuern, C.; Marche, P.N.  
Immunogenetics 30, 137-140, 1989  
A:Title: A rabbit class II MHC gene with strong similarities to HLA-DRA.  
A:Reference number: A45881; MUID:89339606  
A:Accession: A45881  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-255 <LAV>  
A:Cross-references: GB:M28161; NID:g341842; PIDN:AAA31394.1; PID:g529576  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
F:126-191/Domain: immunoglobulin homology <IMM>

Query Match 0.5%; Score 7; DB 2; Length 255;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
Db 80 FEAQAL 86

RESULT 81  
S28473  
rfbH protein - Vibrio cholerae  
C:Species: Vibrio cholerae  
C:Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 08-Oct-1999  
C:Accession: S28473  
R:Manning, P.A.  
submitted to the EMBL Data Library, May 1991  
A:Reference number: S28467  
A:Accession: S28473  
A:Molecule type: DNA  
A:Residues: 1-257 <MAN>  
A:Cross-references: EMBL:X59554; NID:g48381; PIDN:CAA42139.1; PID:g48388  
C:Genetics:  
A:Gene: rfbH

Query Match 0.5%; Score 7; DB 2; Length 257;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 474 SSANFEF 480  
Db 132 SSANFEF 138

RESULT 82  
A38284  
homeotic protein Quox-7 - Japanese quail  
C:Species: Coturnix coturnix japonica (Japanese quail)  
C:Date: 14-Jun-1991 #sequence\_revision 14-Jun-1991 #text\_change 24-Sep-1999  
C:Accession: A38284  
R:Takahashi, Y.; Le Douarin, N.  
Proc. Natl. Acad. Sci. U.S.A. 87, 7482-7486, 1990  
A:Title: cDNA cloning of a quail homeobox gene and its expression in neural crest-der  
A:Reference number: A38284; MUID:91017530  
A:Accession: A38284  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-259 <TAK>  
A:Cross-references: GB:M57611; GB:M37164; NID:g213619; PIDN:AAA63459.1; PID:g213620

C:Superfamily: unassigned homeobox proteins; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
 F:135-191/Domain: homeobox homology <Hox>

Query Match 0.5%; Score 7; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1046 ASLYGTS 1052  
 |||||  
 Db 225 ASLYGTS 231

## RESULT 83

homeotic protein Hox-8 - chicken  
 N:Alternate names: homeodomain protein msx1  
 C:Species: Gallus gallus (chicken)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 17-Nov-2000  
 C:Accession: J50660; B41635; S18675; H45187  
 R:Nohno, T.; Noji, S.; Koyama, E.; Nishikawa, K.; Myokai, F.; Saito, T.; Taniguchi, S.  
 Biochem. Biophys. Res. Commun. 182, 121-128, 1992  
 A:Title: Differential expression of two msh-related homeobox genes Chox-7 and Chox-8 dur  
 A:Reference number: J50659; MUID:92118001  
 A:Accession: J50660  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-259 <NOH>  
 R:Robert, B.; Lyons, G.; Simandl, B.K.; Kuroiwa, A.; Buckingham, M.  
 Genes Dev. 5, 2363-2374, 1991  
 A:Title: The apical ectodermal ridge regulates Hox-7 and Hox-8 gene expression in develo  
 A:Reference number: A41635; MUID:92090717  
 A:Accession: B41635  
 A:Molecule type: mRNA  
 A:Residues: 120-259 <ROB>  
 A:Cross-references: ENBL:X62541; NID:963499; PIDN:CAA44425.1; PID:963500  
 R:Yokouchi, Y.; Ohsugi, K.; Sasaki, H.; Kuroiwa, A.  
 Development 113, 431-444, 1991  
 A:Title: Chicken homeobox gene Msx-1: structure, expression in limb buds and effect of r  
 A:Reference number: S18675; MUID:92146256  
 A:Accession: S18675  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-259 <YOK>  
 A:Cross-references: ENBL:X62097; NID:963615; PIDN:CAA44007.1; PID:963616  
 R:MacKam, S.; Mahon, K.A.  
 Development 112, 791-806, 1991  
 A:Title: Ghox 4.7: a chick homeobox gene expressed primarily in limb buds with limb-type  
 A:Reference number: H45187  
 A:Accession: H45187  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A:Molecule type: mRNA  
 A:Residues: 148-184 <MAC>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:63352)  
 C:Genetics:  
 A:Gene: Hox-8  
 C:Superfamily: unassigned homeobox proteins; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
 F:135-191/Domain: homeobox homology <Hox>

Query Match 0.5%; Score 7; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1046 ASLYGTS 1052  
 |||||  
 Db 225 ASLYGTS 231

## RESULT 84

D86722

hypothetical protein accA [imported] - Lactococcus lactis subsp. lactis (strain IL1  
 C:Species: Lactococcus lactis subsp. lactis  
 C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 23-Mar-2001  
 C:Accession: D86722  
 R:Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.;  
 Genome Res. in press, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium.  
 A:Reference number: A86625  
 A:Accession: D86722  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-260 <STO>  
 A:Cross-references: GB:AE005176; NID:912723699; PIDN:AAK04878.1; GSPDB:GN00146  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: accA

Query Match 0.5%; Score 7; DB 2; Length 260;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AIVGGIA 49  
 |||||  
 Db 57 AIVGGIA 63

## RESULT 85

D82918  
 conserved hypothetical ATP/GTP-binding protein UU217 [imported] - Ureaplasma urealyt  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: D82918  
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
 submitted to GenBank, February 2000  
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of  
 A:Reference number: A82870  
 A:Accession: D82918  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-262 <GLA>  
 A:Cross-references: GB:AE002121; GB:AF222894; NID:96899184; PIDN:AAF30625.1; GSPDB:  
 A:Experimental source: serovar 3; biovar 1  
 C:Genetics:  
 A:Gene: UU217  
 A:Genetic code: SGC3  
 C:Superfamily: conserved hypothetical protein H11714

Query Match 0.5%; Score 7; DB 2; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 623 ARNIKNV 629  
 |||||  
 Db 63 ARNIKNV 69

## RESULT 86

C70881  
 Probable thyA protein - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: C70881  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordo  
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete gen  
 A:Reference number: A70500; MUID:98295987  
 A:Accession: C70881  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-263 <COL>  
A:Cross-references: GB:AL008967; GB:AL123456; NID:g3261491; PIDN:CAA15560.1; PID:g262428  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: thyA  
C:Superfamily: thymidylate synthase; thymidylate synthase homology  
F:1-263/Domain: thymidylate synthase homology <TDS>

Query Match 0.5%; Score 7; DB 2; Length 263;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 590 TGTSLF 596  
|||||  
DB 24 TGTSLF 30

RESULT 87  
A31299  
Chymotrypsin (EC.3.4.21.1) precursor - human  
C:Species: Homo sapiens (man)  
C>Date: 08-Jun-1989 #sequence\_revision 08-Jun-1989 #text\_change 22-Jun-1999  
C:Accession: A31299  
R:Tomita, N.; Izumoto, Y.; Horii, A.; Doi, S.; Yokouchi, H.; Ogawa, M.; Mori, T.; Matsub  
Biochem. Biophys. Res. Commun. 158, 569-575, 1989  
A:Title: Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsinog  
A:Reference number: A31299; MUID:89134264  
A:Accession: A31299  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-263 <TOM>  
A:Cross-references: GB:M24400; NID:g181189; PIDN:AAA52128.1; PID:g181190  
C:Genetics:  
A:Gene: GDB:CTRB1; CTRB  
A:Cross-references: GDB:119820; OMIM:118890  
A:Map position: 16q23.1-16q23.1  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F:34-256/Domain: trypsin homology <TRY>  
F:75.120.213/Active site: His, Asp, Ser #status predicted

Query Match 0.5%; Score 7; DB 2; Length 263;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AKNTPDK 77  
|||||  
DB 166 AKNTPDK 172

RESULT 88  
E82346  
lipopolysaccharide/O-antigen transport protein VC0246 [imported] - Vibrio cholerae (stra  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: E82346  
R:Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; Emolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: E82346  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <HEI>  
A:Cross-references: GB:AE004113; GB:AE003852; NID:g9654648; PIDN:AAF93422.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0246

A:Map position: 1

Query Match 0.5%; Score 7; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 474 SSANFEF 480  
|||||  
DB 139 SSANFEF 145

RESULT 89

H71727  
DNA polymerase III, delta chain (holB) RP172 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C>Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
C:Accession: H71727  
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630; MUID:99039499  
A:Accession: H71727  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-269 <AND>  
A:Cross-references: GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAA14639.1; PID:el34  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: holB; RP172

Query Match 0.5%; Score 7; DB 2; Length 269;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 NAKNISI 197  
|||||  
DB 72 NAKNISI 78

RESULT 90

LNPHLS  
Lectin precursor - flesh fly (Sarcophaga peregrina)  
C:Species: Sarcophaga peregrina  
C>Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 16-Jun-2000  
C:Accession: S07759; S41119; A25736; B25736  
R:Kobayashi, A.; Hirai, H.; Kubo, T.; Ueno, K.; Nakanishi, Y.; Natori, S.  
Biochim. Biophys. Acta 1009, 244-250, 1989  
A:Title: Cloning and in vitro transcription of the Sarcophaga lectin gene.  
A:Reference number: S07759; MUID:90089397  
A:Accession: S07759  
A:Molecule type: DNA  
A:Residues: 1-38 <KOB>  
A:Cross-references: EMBL:X16659; NID:g10271; PIDN:CAA34645.1; PID:g10272  
R:Matsumi, M.; Kobayashi, A.; Kubo, T.; Natori, S.  
Eur. J. Biochem. 219, 449-454, 1994  
A:Title: Purification and characterization of ATBP, a novel protein that binds to A/T  
A:Reference number: S41119; MUID:94139722  
A:Accession: S41119  
A>Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-283 <MAT>  
A:Cross-references: EMBL:D14870; NID:g287440; PIDN:BAA03586.1; PID:g391905  
R:Takahashi, H.; Komano, H.; Kawaguchi, N.; Kitamura, N.; Nakanishi, S.; Natori, S.  
J. Biol. Chem. 260, 12228-12233, 1985  
A:Title: Cloning and sequencing of cDNA of Sarcophaga peregrina humoral lectin induce  
A:Reference number: A25736; MUID:86008294  
A:Accession: A25736  
A:Molecule type: mRNA  
A:Residues: 1-283 <TAK>  
A:Cross-references: GB:M11673; NID:g161264; PIDN:AAA29983.1; PID:g161265  
A:Accession: B25736

A:Molecule type: protein  
 A:Residues: 24-32;164-171;209-214 <TAK2>  
 A:Note: the initiator Met may be either residue 1 or residue 5  
 C:Comment: This lectin is induced in third instar larvae when the body wall is injured  
 C:Genetics:  
 A:Introns: 43/3  
 C:Superfamily: Sarcophaga lectin; C-type lectin homology  
 C:Keywords: hemolymph; lectin  
 F;1-23/Domain: signal sequence #status predicted <SIG>  
 F;24-283/Product: lectin #status experimental <LTN>  
 F;24-157/Domain: C-type lectin homology <LCH>

Query Match 0.5%; Score 7; DB 1; Length 283;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 NNPDNYK 807  
 Db 121 NNPDNYK 127  
 |||||

RESULT 91  
 A84980  
 DNA-directed DNA polymerase (EC 2.7.7.7) [imported] - Buchnera sp. (strain APS)  
 C:Species: Buchnera sp.  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
 C:Accession: A84980  
 R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
 Nature 407, 81-86, 2000  
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.  
 A:Reference number: A84930; MUID:20445173  
 A:Accession: A84980  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-286 <STO>  
 A:Cross-references: GB:AP000398; GSPDB:GN00144  
 A:Experimental source: strain APS  
 C:Genetics:  
 A:Gene: polA; BU431  
 C:Keywords: nucleotidyltransferase

Query Match 0.5%; Score 7; DB 2; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 NAKNISI 197  
 Db 229 NAKNISI 235  
 |||||

RESULT 92  
 A83334  
 hypothetical protein PA2481 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: A83334  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 ; Lory, S.; Olson, M.V.  
 Nature 405, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: A83334  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-291 <STO>  
 A:Cross-references: GB:AE004091; NID:g9948532; PIDN:AG05869.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2481

Query Match 0.5%; Score 7; DB 2; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 APPEGGY 331  
 Db 159 APPEGGY 165  
 |||||

RESULT 93  
 E86474  
 unknown protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: E86474  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al  
 an, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Ki  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: E86474  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-293 <STO>  
 A:Cross-references: GB:AE005172; NID:g11034941; PIDN:AAG27098.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 0.5%; Score 7; DB 2; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VGLVSI 26  
 Db 128 VGLVSI 134  
 |||||

RESULT 94  
 B64649  
 ATP-binding protein - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 02-Feb-2001  
 C:Accession: B64649  
 R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalek, H.G.; Godek, A.; Mc  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathe  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frase  
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: B64649  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-294 <TM>  
 A:Cross-references: GB:AE000611; GB:AE000511; NID:g2314173; PIDN:AAD08077.1; PID:g2  
 C:Superfamily: Bacillus subtilis flagellar probable biosynthesis switch protein ylx  
 C:Keywords: nucleotide binding; P-loop  
 F;35-42/Region: nucleotide-binding motif A (P-loop)

Query Match 0.5%; Score 7; DB 2; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 954 NNIASLE 960



```

Db 217 NNIASLE 223
|||||
RESULT 95
D71938
hypothetical protein jhp0390 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: D71938
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <ARN>
A:Cross-references: GB:AE001473; GB:AE001439; NID:94154910; PIDN:AAD05975.1; PID:9415492
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0390
C:Superfamily: Bacillus subtilis flagellar probable biosynthesis switch protein ylxH

Query Match 0.5%; Score 7; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 954 NNIASLE 960
|||||
Db 217 NNIASLE 223

RESULT 96
A60131
homeotic protein Xhox-7.1 - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 24-Sep-1999
C:Accession: A60131; S14513
R:Su, M.W.; Suzuki, H.R.; Solursh, M.; Ramirez, F.
Development 111, 1179-1187, 1991
A:Title: Progressively restricted expression of a new homeobox-containing gene during xe
A:Reference number: A60131; MUID:91347929
A:Accession: A60131
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 2-295 <SUA>
A:Cross-references: EMBL:X58773
R:Su, M.W.; Suzuki, R.H.; Solursh, M.; Ramirez, F.
submitted to the EMBL Data Library, December 1990
A:Reference number: S14513
A:Accession: S14513
A:Molecule type: mRNA
A:Residues: 1-295 <SUA>
A:Cross-references: EMBL:X58773; NID:964782; PIDN:CAA41574.1; PID:964783
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;171-227/Domain: homeobox homology <HOX>

Query Match 0.5%; Score 7; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1046 ASLYGTS 1052
|||||
Db 261 ASLYGTS 267

RESULT 97
A71425
hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: A71425
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoft, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palmer, K.; Benes, V.; Rechman, S.;
C.; Chaitazis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A:Reference number: A71400; MUID:98121113
A:Accession: A71425
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-298 <BEV>
A:Cross-references: GB:297339; NID:92244901; PID:e327488; PID:92244949
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match 0.5%; Score 7; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 SAGLNII 324
|||||
Db 150 SAGLNII 156

RESULT 98
G64541
cell binding factor 2 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: G64541
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; Mcke
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467
A:Accession: G64541
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-299 <TOM>
A:Cross-references: GB:AE000538; GB:AE000511; NID:92313263; PIDN:AAD07245.1; PID:9231

Query Match 0.5%; Score 7; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGL 23
|||||
Db 8 LALVGL 14

RESULT 99
B71967
Probable peptidyl-prolyl cis-trans isomerase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: B71967
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

```

A:Reference number: A71800; MUID:99120557  
 A:Accession: B71967  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-299 <ARN>  
 A:Cross-references: GB:AE001454; GB:AE001439; NID:g4154666; PIDN:AA05744.1; PID:g415467  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: jhp0161

Query Match 0.5%; Score 7; DB 2; Length 299;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGL 23  
 -|||||||  
 DB 8 LALVGL 14

## RESULT 100

protein F22C12.24 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: A96665  
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Anen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Chin, C.W.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: A96665  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-299 <STO>  
 A:Cross-references: GB:AE005173; NID:g6692104; PIDN:AAF24569.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F22C12.24  
 A:Map position: 1

Query Match 0.5%; Score 7; DB 2; Length 299;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 SLLSSKI 109  
 |||||  
 DB 247 SLLSSKI 253

## RESULT 101

phosphate butyryltransferase (EC 2.3.1.19) - Clostridium acetobutylicum (strain NCIMB 80  
 N:Alternate names: phosphotransbutyrylase  
 C:Species: Clostridium acetobutylicum  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-May-2000  
 C:Accession: JN0794  
 R:Oultram, J.D.; Burr, I.D.; Elmore, M.J.; Minton, N.P.  
 Gene 131, 107-112, 1993  
 A:Title: Cloning and sequence analysis of the genes encoding phosphotransbutyrylase and  
 A:Reference number: PN0619; MUID:93380658  
 A:Accession: JN0794  
 A:Molecule type: DNA  
 A:Residues: 1-302 <OUL>  
 A:Cross-references: GB:L04468; NID:g144890; PIDN:AAAS2080.1; PID:g144892  
 C:Comment: This enzyme is involved in butyrate generation and catalyzes the conversion of  
 C:Genetics:

A:Gene: ptb  
 C:Superfamily: phosphate acetyltransferase  
 C:Keywords: acyltransferase; coenzyme A

Query Match 0.5%; Score 7; DB 1; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 VIDGPFA 391  
 |||||  
 DB 209 VIDGPFA 215

## RESULT 102

T33899  
 hypothetical protein Y39F10C.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T33899  
 R:Leonard, S.; Graves, T.; Wilson, C.  
 submitted to the EMBL Data Library, February 1999  
 A:Description: The sequence of C. elegans cosmid Y39F10C.  
 A:Reference number: Z21434  
 A:Accession: T33899  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-311 <LBO>  
 A:Cross-references: EMBL:AF125452; PIDN:AA12828.1; GSPDB:GN00020; CESP:Y39F10C.1  
 A:Experimental source: strain Bristol N2; clone Y39F10C  
 C:Genetics:  
 A:Gene: CESP:Y39F10C.1  
 A:Map position: 2  
 A:Introns: 51/3; 86/1; 199/1; 272/1

Query Match 0.5%; Score 7; DB 2; Length 311;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 875 TPNLVAI 881  
 |||||  
 DB 256 TPNLVAI 262

## RESULT 103

A83951  
 riboflavin kinase / FAD synthase ribC [imported] - Bacillus halodurans (strain C-125  
 C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 08-Dec-2000  
 C:Accession: A83951  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.;  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans  
 A:Reference number: A83650; MUID:20263314  
 A:Accession: A83951  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-313 <STO>  
 A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA06128.1; GSPDB:  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: ribC  
 C:Superfamily: conserved hypothetical protein HI0963

Query Match 0.5%; Score 7; DB 2; Length 313;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1005 QKFASLE 1011  
 |||||  
 DB 288 QKFASLE 294

```

RESULT 104
B84291
Hypothetical protein Vng1365c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: B84291
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483
A:Accession: B84291
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-318 <STO>
A:Cross-references: GB:AE004437; NID:gl0580874; PIDN:AA019694.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG1365C

Query Match 0.5%; Score 7; DB 2; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LALVGL 23
|||||
Db 112 LALVGL 118

RESULT 105
A43452
acetyl-CoA carboxylase (EC 6.4.1.2) carboxyltransferase alpha chain [validated] - Escher
C:Species: Escherichia coli
C>Date: 04-Mar-1993 #sequence_revision 08-Feb-1996 #text_change 02-Jun-2000
C:Accession: A43452; D28390; A64743
R:Li, S.J.; Cronan Jr., J.E.
J. Biol. Chem. 267, 16841-16847, 1992
A:Title: The genes encoding the two carboxyltransferase subunits of Escherichia coli ace
A:Reference number: A43452; MUID:92380982
A:Accession: A43452
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA; protein
A:Residues: 1-319 <LI1>
A:Cross-references: GB:M96394; NID:gl47321; PIDN:AA070370.1; PID:gl47322
A:Note: sequence extracted from NCBI backbone (NCBIP:111872)
R:Tomasiewicz, H.G.; Mchenry, C.S.
J. Bacteriol. 169, 5735-5744, 1987
A:Title: Sequence analysis of the Escherichia coli dnaE gene.
A:Reference number: A91855; MUID:98058791
A:Accession: D28390
A:Molecule type: DNA
A:Residues: 1-20 <TON>
A:Cross-references: GB:M19334; GB:M18265; GB:M18266; NID:q450760
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: A64743
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-319 <BLAT>
A:Cross-references: GB:AE000127; GB:U00096; NID:gl7863370; PIDN:AA073296.1; PID:gl786382;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: accA
A:Map position: 4.3 min
C:Complex: in E. coli, acetyl-CoA carboxylase is composed of biotin carboxylase (EC 6.3.
otin carboxyl carrier protein (BCCP, homodimer) (PIR:BKEC9)

```

C:Function: <ACC>  
A:Description: EC 6.4.1.2 [validated; MUID:75035569]; the acetyl-CoA carboxylase comp  
A:Pathway: fatty acid biosynthesis  
C:Function: <CTRA>  
A:Description: catalyzes the transfer of the carboxyl group from the carboxyl-biotin  
C:Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain  
C:Keywords: fatty acid biosynthesis; ligase  
F:84-119/Region: acyl-CoA binding #status predicted

Query Match 0.5%; Score 7; DB 1; Length 319;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 AIVGGIA 49  
|||||  
Db 100 AIVGGIA 106

RESULT 106  
C85503  
Hypothetical protein accA [imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: C85503  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: C85503  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-319 <STO>  
A:Cross-references: GB:AE005174; NID:gl2512913; PIDN:AG54487.1; GSPDB:GN00145; UWGP:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: accA  
C:Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

Query Match 0.5%; Score 7; DB 2; Length 319;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 AIVGGIA 49  
|||||  
Db 100 AIVGGIA 106

RESULT 107  
F82100  
acetyl-CoA carboxylase, carboxyl transferase alpha chain VC2244 [imported] - Vibrio c  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: F82100  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: F82100  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-319 <HEI>  
A:Cross-references: GB:AE004296; GB:AE003852; NID:q9656799; PIDN:AAF95388.1; GSPDB:GN  
A:Experimental source: serogroup O1, strain N16961, biotype El Tor  
C:Genetics:  
A:Gene: VC2244  
A:Map position: 1  
C:Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

Query Match 0.5%; Score 7; DB 2; Length 319;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AIVGGIA 49  
 |||||  
 Db 100 AIVGGIA 106

## RESULT 108

S73795  
 hypothetical protein H91\_orf322 - Mycoplasma pneumoniae (strain ATCC 29342)

C:Species: Mycoplasma pneumoniae  
 A:Variety: ATCC 29342  
 C:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
 C:Accession: S73795  
 R:Himmelfreisch, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
 Nucleic Acids Res. 24, 4420-4449, 1996  
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
 A:Reference number: S73327; MUID:97105885

A:Accession: S73795  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-322 <HTM>

A:Cross-references: EMBL:AE000046; GB:U00089; NID:g1674152; PIDN:AAB96117.1; PID:g167415  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 C:Genetics:  
 A:Genetic code: SGC3

Query Match 0.5%; Score 7; DB 2; Length 322;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 967 QTLSLN 973  
 |||||  
 Db 42 QTLSLN 48

## RESULT 109

G69580

acetyl-CoA carboxylase (EC 6.4.1.2), carboxyltransferase alpha chain - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000

C:Accession: G69580  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 y, M.; Ogawa, K.; Ogilwa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, I.; Sato, T.; Scanlon,  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033

A:Accession: G69580  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-325 <KUN>  
 A:Cross-references: GB:299118; GB:AL009126; NID:g2635200; PIDN:CAB14880.1; PID:g2635385  
 A:Experimental source: strain 168

C:Genetics:  
 A:Gene: accA

C:Function:  
 A:Pathway: fatty acid biosynthesis  
 C:Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

C:Keywords: fatty acid biosynthesis; ligase

Query Match 0.5%; Score 7; DB 2; Length 325;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AIVGGIA 49  
 |||||  
 Db 96 AIVGGIA 102

## RESULT 110

E65084

hypothetical protein b2983 - Escherichia coli (strain K-12)

C:Species: Escherichia coli  
 C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 18-Aug-2000  
 C:Accession: E65084  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.  
 A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617

A:Accession: E65084  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-325 <BLAT>

A:Cross-references: GB:AE000381; GB:U00096; NID:g2367181; PIDN:AAC76019.1; PID:g1789  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Superfamily: Escherichia coli hypothetical protein b2983

Query Match 0.5%; Score 7; DB 2; Length 325;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AIVGGIA 49  
 |||||  
 Db 81 AIVGGIA 87

## RESULT 111

B84263

succinoglycan biosynthesis protein [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: B84263  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lask  
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; J  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.  
 A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A8160; MUID:20504483

A:Accession: B84263

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-328 <STO>

A:Cross-references: GB:AE004437; NID:g10580616; PIDN:AAG19470.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: exom

Query Match 0.5%; Score 7; DB 2; Length 328;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 874 ATPNLVA 880  
 |||||  
 Db 318 ATPNLVA 324

## RESULT 112

A83405  
 probable hydroxylase molybdopterin-containing subunit PA1932 [imported] - Pseudomonas aeruginosa  
 C:Species: Pseudomonas aeruginosa  
 C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: A83405  
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: A83405  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-329 <STO>  
 A:Cross-references: GB:AE004619; GB:AE004091; NID:g9947920; PIDN:AAG05320.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA1932

Query Match 0.5%; Score 7; DB 2; Length 329;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;

QY 19 LVGALVS 25  
 |||||  
 DB 66 LVGALVS 72

RESULT 113  
 A26478  
 porphobilinogen synthase (EC 4.2.1.24) - human  
 A:Alternate names: delta-aminolevulinic acid dehydratase  
 C:Species: Homo sapiens (man)  
 C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 18-Jun-1999  
 C:Accession: A26478; I55569; I70288  
 R:Wetmur, J.G.; Bishop, D.F.; Cantelmo, C.; Desnick, R.J.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 7703-7707, 1986  
 A:Title: Human delta-aminolevulinic acid dehydratase: nucleotide sequence of a full-length cDNA  
 A:Reference number: A26478; MUID:87017017  
 A:Accession: A26478  
 A:Molecule type: mRNA  
 A:Residues: 1-330 <NET>  
 A:Cross-references: GB:M13928; NID:g178328; PIDN:AAA51687.1; PID:g178329  
 R:Ishida, N.; Fujita, H.; Fukuda, Y.; Noguchi, T.; Doss, M.; Kappas, A.; Sassa, S.  
 J. Clin. Invest. 89, 1431-1437, 1992  
 A:Title: Cloning and expression of the defective genes from a patient with delta-aminolevulinic acid dehydratase deficiency  
 A:Reference number: I55569; MUID:92235256  
 A:Accession: I55569  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-239, 'W', 241-330 <ISH1>  
 A:Cross-references: GB:S99468; NID:g248838; PIDN:AAC60581.1; PID:g248839  
 A:Experimental source: mutant allele G1  
 A:Accession: I70288  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-273, 'T', 275-330 <ISH2>  
 A:Cross-references: GB:S99471; NID:g248840; PIDN:AAC60582.1; PID:g248841  
 A:Experimental source: mutant allele G2  
 C:Genetics:  
 A:Gene: GDB:ALAD  
 A:Cross-references: GDB:I19665; OMIM:125270  
 A:Map position: 9q32-q34  
 A:Note: defects in this gene can result in acute hepatic porphyria  
 C:Complex: homocytamer

A:Description: catalyzes the condensation of 2 molecules of 5-aminolevulinic acid to produce uroporphyrinogen  
 A:Pathway: porphyrin biosynthesis  
 C:Superfamily: porphobilinogen synthase  
 C:Keywords: carbon-oxygen lyase; homocytamer; hydro-lyase; porphyrin biosynthesis; zinc binding  
 F:122-132/Region: zinc binding #status predicted

F:252/Active site: Lys #status predicted

Query Match 0.5%; Score 7; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;

QY 948 TATTTLN 954  
 |||||  
 DB 21 TATTTLN 27

RESULT 114  
 B65179  
 ribose operon repressor - Escherichia coli  
 C:Species: Escherichia coli  
 C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 20-Aug-1999  
 C:Accession: B65179; A41828  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: B65179  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-330 <BLAT>  
 A:Cross-references: GB:AE000452; GB:U00096; NID:g1790188; PIDN:AAC76776.1; PID:g17901  
 A:Experimental source: strain K-12, substrain MG1655  
 R:Mauzy, C.A.; Hermodson, M.A.  
 Protein Sci. 1, 831-842, 1992  
 A:Title: Structural and functional analyses of the repressor, RbsR, of the ribose operon  
 A:Reference number: A41828; MUID:93278299  
 A:Accession: A41828  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-11, 'L', 14-330 <MAU>  
 A:Cross-references: GB:M13169; NID:g147511; PIDN:AAA51477.1; PID:g147517  
 C:Comment: The start codon, verified by N-terminal sequencing, is an atypical TTG, an C:Comment: The repressor has been shown to bind the operator sequence TCAGCGAACGTTTC  
 C:Genetics:  
 A:Gene: rbsR  
 C:Superfamily: lac repressor  
 C:Keywords: DNA binding; transcription regulation

Query Match 0.5%; Score 7; DB 2; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;

QY 541 LITASTN 547  
 |||||  
 DB 65 LITASTN 71

RESULT 115  
 H86061  
 regulator for rbs operon [imported] - Escherichia coli (strain O157:H7)  
 C:Species: Escherichia coli  
 C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: H86061  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoukis, K.; Apoda  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: H86061  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-330 <STO>  
 A:Cross-references: GB:AE005174; NID:g12518616; PIDN:AAG58956.1; GSPDB:GN00145; UWGP:  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:

A:Gene: rbsR  
C:Superfamily: lac repressor

Query Match 0.5%; Score 7; DB 2; Length 330;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 LITASTN 547  
|||||||  
DB 65 LITASTN 71

## RESULT 116

S74931  
esterase sll0644 - *Synechocystis* sp. (strain PCC 6803)  
N:Alternate names: protein sll0644  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S74931  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
S.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S74931  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-331 <KAN>  
A:Cross-references: EMBL:D80902; GB:AB001339; NID:g1652027; PIDN:BAAL6971.1; PID:g165204  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: probable lipolytic protein ybac

Query Match 0.5%; Score 7; DB 2; Length 331;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 467 GGYALAG 473  
|||||||  
DB 109 GGYALAG 115

## RESULT 117

G96750  
unknown protein F28P22.19 [imported] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: G96750  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719  
A:Accession: G96750  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-331 <STO>  
A:Cross-references: GB:AE005173; NID:g6648166; PIDN:AAF21166.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

Query Match 0.5%; Score 7; DB 2; Length 331;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 HRKINRP 13  
|||||||  
DB 77 HRKINRP 83

## RESULT 118

D83585  
hypothetical protein PA0497 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83585  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic p  
A:Reference number: A82950; MUID:20437337  
A:Accession: D83585  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <STO>  
A:Cross-references: GB:AE004486; GB:AE004091; NID:g9946345; PIDN:AGC03886.1; GSPDB:G  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA0497

Query Match 0.5%; Score 7; DB 2; Length 333;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1112 QGALGSD 1118  
|||||||  
DB 87 QGALGSD 93

## RESULT 119

RDECEP  
N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) - *Escherichia coli*  
C:Species: *Escherichia coli*  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 11-Jun-1999  
C:Accession: JT0332; A42377; A65203; A30776  
R:Parson, C.; Boyen, A.; Cohen, G.N.; Glansdorff, N.  
Gene 68, 275-283, 1988  
A:Title: Nucleotide sequence of *Escherichia coli* argB and argC genes: comparison of  
es.  
A:Reference number: JT0331; MUID:89121510  
A:Accession: JT0332  
A:Molecule type: DNA  
A:Residues: 1-334 <PAR>  
A:Cross-references: GB:M21446; NID:g145332; PIDN:AAA23477.1; PID:g145333  
R.Meinel, T.; Schmitt, E.; Mechulam, Y.; Blanquet, S.  
J. Bacteriol. 174, 2323-2331, 1992  
A:Title: Structural and biochemical characterization of the *Escherichia coli* argE ge  
A:Reference number: A42377; MUID:92202162  
A:Accession: A42377  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-19 <MEI>  
A:Cross-references: GB:X55417  
R.Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: A65203  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-334 <BLAT>  
A:Cross-references: GB:AE000470; GB:U00096; NID:g2367332; PIDN:AACT6940.1; PID:g1790;

A:Experimental source: strain K-12, substrain MG1655  
 C:Comment: In arginine biosynthesis glutamate is first converted to N-acetylglutamate, which is then  
 e catalyzed by acetylglutamate kinase and N-acetyl-gamma-glutamyl-phosphate reductase.  
 C:Genetics:

A:Gene: argC  
 A:Map position: 90 min  
 C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase  
 C:Keywords: arginine biosynthesis; oxidoreductase  
 F:154/Active site: Cys #status predicted

Query Match 0.5%; Score 7; DB 1; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKA 215  
 |||||  
 Db 187 SGAGRKA 193

## RESULT 120

D86087  
 N-acetyl-gamma-glutamylphosphate reductase [imported] - Escherichia coli (strain O157:H7)  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001

C:Accession: D86087  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D86087  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-334 <STO>  
 A:Cross-references: GB:AE005174; NID:q12518880; PIDN:AAG59160.1; GSPDB:GN00145; UWGP:Z55  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:

A:Gene: argC  
 C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase

Query Match 0.5%; Score 7; DB 2; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKA 215  
 |||||  
 Db 187 SGAGRKA 193

## RESULT 121

D83582  
 hypothetical protein PA0498 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: D83582  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: D83582  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-335 <STO>

A:Cross-references: GB:AE004487; GB:AE004091; NID:g9946361; PIDN:AAG03887.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:

A:Gene: PA0498

Query Match 0.5%; Score 7; DB 2; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1112 QGALGSD 1118  
 |||||  
 Db 88 QGALGSD 94

## RESULT 122

T51045  
 hypothetical protein B15120.150 [imported] - Neurospora crassa

C:Species: Neurospora crassa  
 C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 08-Sep-2000  
 C:Accession: T51045  
 R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu  
 submitted to the Protein Sequence Database, July 2000

A:Reference number: 225286  
 A:Accession: T51045  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-340 <SCH>  
 A:Cross-references: EMBL:AL389900; GSPDB:GN00116; NCSP:B15120.150  
 A:Experimental source: BAC clone B15120; strain OR74A  
 C:Genetics:  
 A:Gene: NCSP:B15120.150  
 A:Map position: 6  
 A:Introns: 158/1; 209/3  
 C:Superfamily: Neurospora crassa hypothetical protein B15120.150

Query Match 0.5%; Score 7; DB 2; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 588 LETGTRS 594  
 |||||  
 Db 72 LETGTRS 78

## RESULT 123

S57510  
 Replication initiation protein - Selenomonas ruminantium

C:Species: Selenomonas ruminantium  
 C:Date: 10-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 24-Sep-1999  
 C:Accession: S57510  
 R:Murray, J.J.J.; Hazlewood, G.G.P.; Gilbert, H.H.J.

submitted to the EMBL Data Library, June 1995  
 A:Description: Sequencing and characterization of a RC plasmid from the rumen bacteri  
 A:Reference number: S57510

A:Accession: S57510  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-342 <MUR>  
 A:Cross-references: EMBL:Z49917; NID:9872307; PIDN:CAA90151.1; PID:9872308  
 C:Superfamily: replication protein

Query Match 0.5%; Score 7; DB 2; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 998 RLQALKD 1004  
 |||||  
 Db 66 RLQALKD 72

## RESULT 124

C64978

hypothetical 36.9 kD protein in mrp 5' region precursor - Escherichia coli (strain K-1  
 C:Species: Escherichia coli  
 C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
 C:Accession: C64978

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of *Escherichia coli* K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: C64978  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-344 <BLAT>  
 A:Cross-references: GB:AB000300; GB:U00096; NID:g1788425; PIDN:AAC75169.1; PID:g1788426;  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: yehA

Query Match 0.5%; Score 7; DB 2; Length 344;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 861 NALQNA 867  
 |||||  
 Db 172 NALQNA 178

RESULT 125  
 A44164  
 secreted glycoprotein, 43 K - trichina  
 C:Species: *Trichinella spiralis* (trichina)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Sep-1997  
 C:Accession: A44164; A60630; S27862  
 R:Vasiliadis, D.K.; Despommier, D.; Misek, D.E.; Polvere, R.I.; Gold, A.M.; Van der Ploeg  
 J. Biol. Chem. 267, 18459-18465, 1992  
 A:Title: Analysis of a 43-kDa glycoprotein from the intracellular parasitic nematode *Trichinella*  
 A:Reference number: A44164; MUID:92406752  
 A:Accession: A44164  
 A:Molecule type: mRNA  
 A:Residues: 1-344 <VAS>  
 A:Cross-references: EMBL:M95499; NID:g162534; PID:g162535  
 A:Note: sequence extracted from NCBI backbone (NCBIN:113310, NCBIPI:113321)  
 A:Note: part of this sequence was confirmed by protein sequencing  
 R:Gold, A.M.; Despommier, D.D.; Buck, S.W.  
 Mol. Biochem. Parasitol. 41, 187-196, 1990  
 A:Title: Partial characterization of two antigens secreted by L1 larvae of *Trichinella*  
 A:Reference number: A60630; MUID:90377287  
 A:Accession: A60630  
 A:Molecule type: protein  
 A:Residues: 23, 'X', 25-62, 'YGSF', 66-80, 'T'; 115-116, 'X', 118-132; 249-278 <GOL>  
 C:Keywords: glycoprotein

Query Match 0.5%; Score 7; DB 2; Length 344;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 736 LKAKII 742  
 |||||  
 Db 45 LKAKII 51

RESULT 126  
 S72490  
 N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) - *Bacillus stearothermophilus*  
 C:Species: *Bacillus stearothermophilus*  
 C:Date: 24-Oct-1998 #sequence\_revision 24-Oct-1998 #text\_change 11-Jun-1999  
 C:Accession: S72490; I39765  
 R:Savchenko, A.; Charlier, D.; Dion, M.; Weigel, P.; Hallet, J.N.; Holtham, C.; Baumberg  
 Mol. Gen. Genet. 252, 69-78, 1996  
 A:Title: The arginine operon of *Bacillus stearothermophilus*: characterization of the con  
 A:Reference number: S72490; MUID:96397511  
 A:Accession: S72490  
 A>Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-84 <SAV>

A:Experimental source: strain NCIB8224  
 R:Sakanyan, V.; Charlier, D.; Legrain, C.; Kochikyan, A.; Mett, I.; Pierard, P.; Gla  
 J. Gen. Microbiol. 139, 393-402, 1993  
 A:Title: Primary structure, partial purification and regulation of key enzymes of th  
 e.  
 A:Reference number: I39765; MUID:93232760  
 A:Accession: I39765  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 52-345 <RES>  
 A:Cross-references: GB:L06036; NID:g304133; PIDN:AAA22196.1; PID:g304134  
 A:Experimental source: strain NCIB8224  
 C:Genetics:  
 A:Gene: argC  
 C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase  
 C:Keywords: arginine biosynthesis; oxidoreductase  
 F:149/Active site: Cys #status predicted

Query Match 0.5%; Score 7; DB 2; Length 345;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKA 215  
 |||||  
 Db 181 SGAGRKA 187

RESULT 127  
 E71843  
 oxygen-independent coproporphyrinogen III oxidase - *Helicobacter pylori* (strain J99)  
 C:Species: *Helicobacter pylori*  
 A:Variety: strain J99  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 11-Jun-1999  
 C:Accession: E71843  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, G  
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G  
 Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric  
 A:Reference number: A71800; MUID:99120557  
 A:Accession: E71843  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-345 <ARN>  
 A:Cross-references: GB:AE001542; GB:AE001439; NID:g4155739; PIDN:AAD06720.1; PID:g41  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: hemN\_2  
 C:Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 0.5%; Score 7; DB 2; Length 345;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1260 ELKLAKE 1266  
 |||||  
 Db 170 ELKLAKE 176

RESULT 128  
 C82052  
 N-acetyl-gamma-glutamyl-phosphate reductase VC3644 [imported] - *Vibrio cholerae* (str  
 C:Species: *Vibrio cholerae*  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: C82052  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R  
 ; Carlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Seller  
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A:Reference number: A82035; MUID:20406833  
 A:Accession: C82052



A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-346 <HEI>  
A:Cross-references: GB:AE004330; GB:AE003852; NID:g9657225; PIDN:AAF95785.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:

A:Map position: 1  
A:Gene: VC2644  
C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase

Query Match 0.5%; Score 7; DB 2; Length 346;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKA 215  
DB 200 SGAGRKA 206  
|||||

# RESULT 129

probable ribonucleoside-diphosphate reductase, subunit beta (nrdb) - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
C:Accession: G71372

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
son, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo  
nald, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; NUID:98332770

A:Accession: G71372

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-351 <COL>

A:Cross-references: GB:AE001190; GB:AE000520; NID:g3322303; PIDN:AAC65049.1; PID:g332231

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0053

C:Superfamily: ribonucleoside diphosphate reductase beta chain

Query Match 0.5%; Score 7; DB 2; Length 351;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 842 GGNNTNL 848  
DB 29 GGNNTNL 35  
|||||

# RESULT 130

hypothetical protein NMA1323 [imported] - Neisseria meningitidis (strain Z491 serogroup  
C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001

C:Accession: C81901

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z491.

A:Reference number: A81775; NUID:20222556

A:Accession: C81901

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-351 <PAR>

A:Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84574.1; PID:g737999

A:Experimental source: serogroup A, strain Z491

C:Genetics:

A:Gene: NMA1323

Query Match 0.5%; Score 7; DB 2; Length 351;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 NALVLKP 1164  
DB 253 NALVLKP 259  
|||||

# RESULT 131

coporophyrinogen oxidase (EC 1.3.3.3) III, oxygen-independent - Helicobacter pylori  
C:Species: Helicobacter pylori  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: B64673

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,  
Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: B64673

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-352 <TOM>

A:Cross-references: GB:AE000628; GB:AE000511; NID:g2314386; PIDN:AAD08271.1; PID:g231

C:Genetics:

A:Start codon: GTG

C:Superfamily: oxygen-independent coproporphyrinogen oxidase

C:Keywords: oxidoreductase

Query Match 0.5%; Score 7; DB 1; Length 352;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1260 ELKAKE 1266  
DB 177 ELKAKE 183  
|||||

# RESULT 132

phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) - Helicobacter pylori  
C:Species: Helicobacter pylori  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: E64581

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,  
Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: E64581

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-353 <TOM>

A:Cross-references: GB:AE000564; GB:AE000511; NID:g2313602; PIDN:AAD07559.1; PID:g231

C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase

C:Keywords: transferase

Query Match 0.5%; Score 7; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VSLALVG 21  
DB 235 VSLALVG 241  
|||||

## RESULT 133

C39725

hypothetical protein (MS11 5' region) - fruit fly (*Drosophila melanogaster*)C:Species: *Drosophila melanogaster*

C&gt;Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 16-Feb-1997

C:Accession: C39725

R:Samuels, M.E.; Schedl, P.; Cline, T.W.

Mol. Cell. Biol. 11, 3584-3602, 1991

A:Title: The complex set of late transcripts from the *Drosophila* sex determination gene

A:Reference number: A39725; MUID:91260708

A:Accession: C39725

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-355 &lt;SAM&gt;

A:Cross-references: GB:M59488

C:Genetics:

A:Gene: FlyBase:Sc1

A:Cross-references: FlyBase:FBgn0003659

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 355;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 GHFASNL 1290

Db 206 GHFASNL 212

## RESULT 134

GB4904

probable DOF zinc finger protein [imported] - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C&gt;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: GB4904

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487

A:Accession: GB4904

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-356 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:g4415939; PIDN:AAD20169.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g46590

A:Map position: 2

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 356;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 851 NNTSNAR 857

Db 20 NNTSNAR 26

## RESULT 135

T45934

hypothetical protein F5K20.240 - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C&gt;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 17-Mar-2000

C:Accession: T45934

R:Monfort, A.; Casacuberta, E.; Pulgomech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223017

A:Accession: T45934

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-358 &lt;MON&gt;

A:Cross-references: EMBL:AL132960

A:Experimental source: cultivar Columbia; BAC clone F5K20

C:Genetics:

A:Map position: 3

A:Introns: 98/3; 157/3; 207/3; 234/3; 258/2; 281/1

A:Note: F5K20.240

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 358;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 946 LNTATTT 952

Db 16 LNTATTT 22

## RESULT 136

E84890

Probable dihydronicotinate synthase [imported] - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C&gt;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: E84890

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487

A:Accession: E84890

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-365 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:g2583111; PIDN:AAB82620.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g45440

A:Map position: 2

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 365;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 NTDDIKA 789

Db 59 NTDDIKA 65

## RESULT 137

T16913

hypothetical protein T20H4.4 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C&gt;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16913

R:Du, Z.

submitted to the EMBL Data Library, February 1994

A:Description: The sequence of *C. elegans* cosmid T20H4.

A:Reference number: 218603

A:Accession: T16913

A:Status: preliminary; translated from GB/EMBL/DDB3

A:Molecule type: DNA

A:Residues: 1-367 &lt;DUZ&gt;

A:Cross-references: EMBL:U00037; NID:g459006; PID:g459010; PIDN:AAA50661.1; CESP:T20H4

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:T20H4.4

A:Introns: 122/3; 200/3; 310/3

## Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 367;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 NALVLKP 1164  
 Db 97 NALVLKP 103

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 138  
 D70076  
 ethanolamine transporter homolog yxeR - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
 C:Accession: D70076  
 R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
 C.; Bron, S.; Broillett, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F  
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
 Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
 T.; Winters, P.; Wipat, A.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033  
 A:Accession: D70076  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-370 <HIM>  
 A:Cross-references: GB:299124; GB:AL009126; NID:g2636442; PIDN:CAB15981.1; PID:ell84670;  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yxeR

Query Match 0.5%; Score 7; DB 2; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 467 GGYALAG 473  
 Db 90 GGYALAG 96

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 139  
 S73941  
 oligopeptide transport system permease protein amid - Mycoplasma pneumoniae (strain ATCC  
 N:Alternate names: Hypothetical protein G07\_Orf376  
 C:Species: Mycoplasma pneumoniae  
 A:Variety: ATCC 29342  
 C>Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 11-Jan-2000  
 C:Accession: S73941  
 R:Himmelfreich, R.; Hilbert, H.; Plagens, H.; Pirkil, E.; Li, B.C.; Herrmann, R.  
 Nucleic Acids Res. 24, 4420-4449, 1996  
 A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
 A:Reference number: S73327; MUID:97105885  
 A:Accession: S73941  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-376 <HIM>  
 A:Cross-references: EMBL:AE000058; GB:U00089; NID:g1674291; PIDN:AAB96263.1; PID:g167431  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 C:Genetics:  
 A:Gene: amid  
 C:Superfamily: oligopeptide permease protein oppB

Query Match 0.5%; Score 7; DB 2; Length 376;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;

QY 860 NNALQON 866  
 Db 78 NNALQON 84

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 140  
 T52651  
 probable topoisomerase VIA [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 24-Oct-2000  
 C:Accession: T52651  
 R:Hartung, F.; Puchta, H.  
 Nucleic Acids Res. 28, 1548-1554, 2000  
 A>Title: Molecular characterization of two paralogous SPO11 homologues in Arabidopsis  
 A:Reference number: 226158  
 A:Accession: T52651  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-383 <HAR>  
 A:Cross-references: EMBL:AJ251990; PIDN:CAB81545.1  
 A:Experimental source: cultivar Columbia; flowers  
 C:Genetics:  
 A>Note: SPO11

Query Match 0.5%; Score 7; DB 2; Length 383;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 SLLSSKI 109  
 Db 331 SLLSSKI 337

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 141  
 E64707  
 outer membrane protein - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 C:Accession: E64707  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
 A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: E64707  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-388 <TOM>  
 A:Cross-references: GB:AE000648; GB:AE000511; NID:g2314670; PIDN:AA08546.1; PID:g231  
 C:Genetics:  
 A:Start codon: GTG

Query Match 0.5%; Score 7; DB 2; Length 388;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 405 TNADGTI 411  
 Db 267 TNADGTI 273

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 142  
 C71813  
 probable outer membrane protein - Helicobacter pylori (strain J99)  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99  
 C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999

C;Accession: C71813  
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
 Nature 397, 176-180, 1999  
 A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
 A;Reference number: A71800; MUID:99120557  
 A;Accession: C71813  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-388 <ARN>  
 A;Cross-references: GB:AE001561; GB:AE001439; NID:94156000; PIDN:AAD06975.1; PID:9415601  
 A;Experimental source: strain J99  
 C;Genetics:  
 A;Gene: jhpl394

Query Match 0.5%; Score 7; DB 2; Length 388;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 405 TNADGTI 411  
 |||||  
 DB 267 TNADGTI 273

## RESULT 143

G70810  
 Probable transcription regulator Rv0823c - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C;Accession: G70810  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987  
 A;Accession: G70810  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-389 <COL>  
 A;Cross-references: GB:AL022004; GB:AL123456; NID:93261550; PIDN:CAAL17629.1; PID:e125396  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: Rv0823c

Query Match 0.5%; Score 7; DB 2; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 VGTVSGL 60  
 |||||  
 DB 49 VGTVSGL 55

## RESULT 144

I39570  
 Probable porin [similarity] - Alcaligenes eutrophus  
 N;Alternate names: protein 9 gbd-region  
 C;Species: Alcaligenes eutrophus  
 C;Date: 09-Mar-1996 #sequence\_revision 09-Mar-1996 #text\_change 09-Jun-2000  
 C;Accession: I39570  
 R;Valentin, H.E.; Zwillingmann, G.; Schonebaum, A.; Steinbuechel, A.  
 Eur. J. Biochem. 227, 43-60, 1995  
 A;Title: Metabolic pathway for biosynthesis of poly(3-hydroxybutyrate-co-4-hydroxybutyrate)  
 A;Reference number: I39561; MUID:95154322  
 A;Accession: I39570  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-390 <RES>  
 A;Cross-references: GB:J36817; NID:9695273; PIDN:AAC41427.1; PID:g695281

C;Superfamily: Bordetella outer membrane protein ompQ

Query Match 0.5%; Score 7; DB 2; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 489 GTATFNN 495  
 |||||  
 DB 290 GTATFNN 296

## RESULT 145

T30149  
 Hypothetical protein C39H7.4 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T30149  
 R;Le, T.T.  
 Submitted to the EMBL Data Library, May 1996  
 A;Description: The sequence of C. elegans cosmid C39H7.  
 A;Reference number: Z20744  
 A;Accession: T30149  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-391 <LET>  
 A;Cross-references: EMBL:U58754; PIDN:AAB00671.1; GSPDB:GN00022; CESP:C39H7.4  
 A;Experimental source: strain Bristol N2; Clone C39H7  
 C;Genetics:  
 A;Gene: CESP:C39H7.4  
 A;Map position: 4  
 A;Introns: 39/1; 275/3; 308/2; 365/1

Query Match 0.5%; Score 7; DB 2; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 RLALYNN 772  
 |||||  
 DB 192 RLALYNN 198

## RESULT 146

T29412  
 Hypothetical protein F21F8.3 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
 C;Accession: T29412  
 R;Wilson, R.; Favello, A.; Le, T.T.  
 Submitted to the EMBL Data Library, April 1997  
 A;Description: The sequence of C. elegans cosmid F21F8.  
 A;Reference number: Z20618  
 A;Accession: T29412  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-393 <WIL>  
 A;Cross-references: EMBL:U97000; PIDN:AAC47991.1; GSPDB:GN00023; CESP:F21F8.3  
 A;Experimental source: strain Bristol N2; clone F21F8  
 C;Genetics:  
 A;Gene: CESP:F21F8.3  
 A;Map position: 5  
 A;Introns: 105/2; 176/3; 302/3

Query Match 0.5%; Score 7; DB 2; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 TNIGTLD 314  
 |||||  
 DB 316 TNIGTLD 322

## RESULT 147

B64709  
transcription termination factor nusa - Helicobacter pylori (strain 26695)  
C/Species: Helicobacter pylori  
C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 15-Oct-1999  
C/Accession: B64709  
R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khakhria, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997  
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A/Reference number: A64520; MUID:97394467  
A/Accession: B64709  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-395 <TOM>  
A/Cross-references: GB:AE000649; GB:AE000511; NID:g2314687; PIDN:RAD08555.1; PID:g231469  
C/Species: Bacillus transcription termination factor nusa; transcription termination  
C/Keywords: transcription termination  
F:56-364/Domain: transcription termination factor nusa homology <TTN>

Query Match 0.5%; Score 7; DB 2; Length 395;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 ICKGGVN 435  
Db 343 ICKGGVN 349  
|||||

## RESULT 148

G71811  
transcription termination factor nusa - Helicobacter pylori (strain J99)  
N/Alternate names: N utilization substance protein A (nusa)  
C/Species: Helicobacter pylori  
A/Variety: strain J99  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 15-Oct-1999  
C/Accession: G71811  
R/Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F. Nature 397, 176-180, 1999  
A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A/Reference number: A71800; MUID:99120557  
A/Accession: G71811  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-395 <ARN>  
A/Cross-references: GB:AE001562; GB:AE001439; NID:g4156017; PIDN:AAD06980.1; PID:g415602  
A/Experimental source: strain J99  
C/Genetics:  
A/Gene: nusa  
C/Superfamily: Bacillus transcription termination factor nusa; transcription termination  
C/Keywords: transcription termination  
F:66-364/Domain: transcription termination factor nusa homology <TTN>

Query Match 0.5%; Score 7; DB 2; Length 395;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 ICKGGVN 435  
Db 343 ICKGGVN 349  
|||||

## RESULT 149

D69378  
3-ketoacyl-CoA thiolase (fdaA-1) homolog - Archaeoglobus fulgidus  
C/Species: Archaeoglobus fulgidus

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 18-Jun-1999  
C/Accession: D69378  
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E. Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997  
A/Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, Smith, H.O.; Woese, C.R.; Venter, J.C.  
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
A/Reference number: A69250; MUID:98049343  
A/Accession: D69378  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-396 <KLE>  
A/Cross-references: GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB90215.1; PID:g264  
C/Superfamily: acetyl-CoA acetyltransferase

Query Match 0.5%; Score 7; DB 2; Length 396;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 739 AKIIGVG 745  
Db 277 AKIIGVG 283  
|||||

## RESULT 150

T35713  
probable oxidoreductase - Streptomyces coelicolor  
C/Species: Streptomyces coelicolor  
C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 19-May-2000  
C/Accession: T35713  
R/Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, January 1998  
A/Reference number: Z21548  
A/Accession: T35713  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-397 <MUR>  
A/Cross-references: EMBL:AL021411; PIDN:CAA16205.1; GSPDB:GN00070; SCOEDB:SC7H1.18  
A/Experimental source: strain A3(2)  
C/Genetics:  
A/Gene: SCOEDB:SC7H1.18  
C/Superfamily: tetracycline 6-hydroxylase

Query Match 0.5%; Score 7; DB 2; Length 397;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 309 SLALVGA 315  
|||||

## RESULT 151

T00658  
hypothetical protein F316.21 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 22-Oct-1999  
C/Accession: T00658  
R/Federspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Arau  
; Vysotskaia, V.S.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W. submitted to the EMBL Data Library, February 1998  
A/Reference number: Z14197  
A/Accession: T00658  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-401 <FED>  
A/Cross-references: EMBL:AC002396; NID:g2749918; PID:g2829881; GSPDB:GN00059; ATSP:F3  
C/Genetics:  
A/Gene: ATSP:F316.21

A:Map position: 1  
A:Introns: 197/3; 236/3; 262/3; 288/1; 309/3; 339/3; 364/3; 375/3

Query Match 0.5%; Score 7; DB 2; Length 401;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LSSKIDG 111  
|||||  
Db 315 LSSKIDG 321

RESULT 152  
VCBBBL  
coat protein precursor - boolarra virus  
N:Contains: coat protein alpha; coat protein beta; coat protein gamma  
C:Species: boolarra virus  
C>Date: 30-Jun-1990 #sequence\_revision 30-Jun-1990 #text\_change 23-Jul-1999  
C:Accession: A34011; S11038  
R:Dasgupta, R.; Sgro, J.Y.  
Nucleic Acids Res. 17, 7525-7526, 1989  
A:Title: Nucleotide sequences of three nodavirus RNA2's: the messenger for their coat pr  
A:Reference number: A34011; MUID:90016821  
A:Accession: A34011  
A:Molecule type: genomic RNA  
A:Residues: 1-403 <DAS>  
A:Cross-references: EMBL:X15960; NID:g58723; PIDN:CAA34082.1; PID:g58724  
R:Kaesberg, P.; Dasgupta, R.; Sgro, J.Y.; Wery, J.P.; Selling, B.H.; Hosur, M.V.; Johnsc  
J. Mol. Biol. 214, 423-435, 1990  
A:Title: Structural homology among four nodaviruses as deduced by sequencing and X-ray c  
A:Reference number: S11036; MUID:90339486  
A:Contents: annotation  
A:Genetics:  
C:Map position: segment 2  
C:Superfamily: black beetle virus coat protein  
C:Keywords: coat protein; RNA binding  
F:1-403/Product: coat protein alpha #status predicted <ALP>  
F:1-359/Product: coat protein beta #status predicted <BET>  
F:25-34/Region: arginine-rich RNA-binding pattern  
F:360-403/Product: coat protein gamma #status predicted <GAM>  
F:359-360/Cleavage site: Asn-Ala (autolytic) #status predicted

Query Match 0.5%; Score 7; DB 1; Length 403;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 VGGIATG 51  
|||||  
Db 384 VGGIATG 390

RESULT 153  
S40976  
hypothetical protein F58A4.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 13-Sep-1998  
C:Accession: S40976  
R:Berkas, M.  
submitted to the EMBL Data Library, February 1992  
A:Reference number: S40973  
A:Accession: S40976  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-410 <BER>  
A:Cross-references: EMBL:222179; NID:g297956; PID:g297959  
A:Introns: 34/1; 120/2; 190/3; 292/3; 321/1; 360/2  
C:Superfamily: DNA primase 50K chain

Query Match 0.5%; Score 7; DB 2; Length 410;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 746 NVSTGTN 752  
|||||  
Db 317 NVSTGTN 323

RESULT 154  
B35670  
protein-tyrosine kinase (EC 2.7.1.112) 2 - slime mold (Dictyostelium discoideum) (f  
C:Species: Dictyostelium discoideum  
C>Date: 28-Sep-1990 #sequence\_revision 28-Sep-1990 #text\_change 04-Feb-2000  
C:Accession: B35670  
R:Tan, J.L.; Spudich, J.A.  
Mol. Cell. Biol. 10, 3578-3583, 1990  
A:Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium d  
A:Reference number: A35670; MUID:90287147  
A:Accession: B35670  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-410 <TAN>  
A:Cross-references: GB:M33784; NID:gl67777; PIDN:AAA33203.1; PID:gl67778  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase  
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase  
F:106-371/Domain: protein kinase homology <KIN>  
F:114-122/Region: protein kinase ATP-binding motif

Query Match 0.5%; Score 7; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 EEOFKER 766  
|||||  
Db 143 EEOFKER 149

RESULT 155  
T13255  
hypothetical protein - Lactococcus lactis phage BK5-T (fragment)  
C:Species: Lactococcus lactis phage BK5-T  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 13-Aug-1999  
C:Accession: T13255  
R:Boyce, J.D.; Davidson, B.E.; Hillier, A.J.  
Appl. Environ. Microbiol. 61, 4089-4098, 1995  
A:Title: Sequence analysis of the temperate Lactococcus lactis bacteriophage BK5-T  
A:Reference number: Z17646; MUID:96064422  
A:Accession: T13255  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-410 <BOY>  
A:Cross-references: EMBL:L44593; NID:g928826; PID:g928827; PIDN:AAA98578.1

Query Match 0.5%; Score 7; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 INNQGTI 683  
|||||  
Db 24 INNQGTI 30

RESULT 156  
A34170  
acrosin (EC 3.4.21.10) precursor - pig  
N:Alternate names: 53K fucose-binding protein  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A34170; S08994; S02428; S04940; S16557; S02780; S10695; S12968  
R:Baba, T.; Kashiwabara, S.; Watanabe, K.; Itoh, H.; Michikawa, Y.; Kimura, K.; Tak  
J. Biol. Chem. 264, 11920-11927, 1989

A:Title: Activation and maturation mechanisms of boar acrosin zymogen based on the deduced  
 A:Reference number: A34170; MUID:89308595  
 A:Accession: A34170  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-415 <BAB>  
 A:Cross-References: GB:J04950; NID:q164702; PIDN:AAA31131.1; PID:q164703  
 R:Ceckova, D.; Toepfer-Petersen, E.; Zucker, A.; Jonakova, V.  
 Biol. Chem. Hoppe-Seyler 371, 317-323, 1990  
 A:Title: Is sperminogen a modified proacrosin? Isolation, purification, and partial char  
 A:Reference number: S08994; MUID:90253655  
 A:Accession: S08994  
 A:Molecule type: protein  
 A:Residues: 'X', 18, 'X', 20-25, 'X', 27-32, 'X', 34-38, 'X', 40-50 <CEC>  
 R:Toepfer-Petersen, E.; Henschen, A.  
 FEBS Lett. 226, 38-42, 1987  
 A:Title: Acrosin shows zona and fucose binding, novel properties for a serine proteinase  
 A:Reference number: S02428; MUID:88083633  
 A:Accession: S02428  
 A:Molecule type: protein  
 A:Residues: 17-32; 40-55 <TOE>  
 R:Adham, I.M.; Klemm, U.; Maier, W.M.; Hoyer-Fender, S.; Tsousidou, S.; Engel, W.  
 Eur. J. Biochem. 182, 563-568, 1989  
 A:Title: Molecular cloning of preproacrosin and analysis of its expression pattern in sp  
 A:Reference number: S04940; MUID:89325301  
 A:Accession: S04940  
 A:Molecule type: mRNA  
 A:Residues: 1-7, 9-210, 'Q', 212-216, 'VT', 219-346, 'A', 348-388, 390-393, 'GN', 396, 'LVE', 399-40  
 A:Cross-References: EMBL:X14844  
 A:Note: The authors translated the codon CCT for residue 240 as Ala, GCC for residue 264  
 R:Adham, I.M.  
 submitted to the EMBL Data Library, March 1989  
 A:Reference number: S16657  
 A:Accession: S16657  
 A:Molecule type: mRNA  
 A:Residues: 1-7, 9-210, 'Q', 212-216, 'VT', 219-346, 'A', 348-388, 390-398, 'KELL' <AD2>  
 A:Cross-References: EMBL:X14844; NID:q1867; PIDN:CAA32948.1; PID:q1868  
 A:Note: The difference at the carboxyl end is due to a frameshift error  
 R:Baba, T.; Michikawa, Y.; Kawakura, K.; Arai, Y.  
 FEBS Lett. 244, 132-136, 1989  
 A:Title: Activation of boar proacrosin is effected by processing at both N- and C-termin  
 A:Reference number: S02780; MUID:89171246  
 A:Accession: S02780  
 A:Molecule type: protein  
 A:Residues: 17-69 <BA2>  
 R:Toepfer-Petersen, E.; Steinberger, M.; von Eschenbach, C.E.; Zucker, A.  
 FEBS Lett. 265, 51-54, 1990  
 A:Title: Zona pellucida-binding of boar sperm acrosin is associated with the N-terminal  
 A:Reference number: S10695; MUID:90306316  
 A:Accession: S10695  
 A:Molecule type: protein  
 A:Residues: 40-62 <TO2>  
 R:Toepfer-Petersen, E.; Calvete, J.; Schaefer, W.; Henschen, A.  
 FEBS Lett. 275, 139-142, 1990  
 A:Title: Complete localization of the disulfide bridges and glycosylation sites in boar  
 A:Reference number: S12968; MUID:91085546  
 A:Accession: S12968  
 A:Molecule type: protein  
 A:Residues: 17-29; 34-66; 88-91; 94-121; 123-166; 171-184; 190-207; 209-216; 219-228; 231-245; 248  
 C:Superfamily: acrosin; trypsin homology  
 C:Keywords: glycoprotein; hydrolase; serine proteinase; sperm  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-415/Product: acrosin #status experimental <MAY>  
 F:17-39/Product: acrosin light (A) chain #status experimental <LCH>  
 F:40-415/Product: acrosin heavy (B) chain #status experimental <HCH>  
 F:40-283/Domain: trypsin homology <TRY>  
 F:300-374/Region: proline-rich  
 F:19, 208/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:22-152, 26-160, 71-87, 175-244, 207-223, 234-264/Disulfide bonds: #status experimental  
 F:86, 140, 238/Active site: His, Asp, Ser #status predicted

Query Match

0.5%; Score 7; DB 1; Length 415;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 994 SFAKRLQ 1000  
 |||||  
 DB 378 SFAKRLQ 384  
 RESULT 157  
 S11674  
 acrosin (EC 3.4.21.10) precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
 C:Accession: S11674; S23499; S12063; A61022; S03330  
 R:Keime, S.; Adham, I.M.; Engel, W.  
 Eur. J. Biochem. 190, 195-200, 1990  
 A:Title: Nucleotide sequence and exon-intron organization of the human proacrosin gen  
 A:Reference number: S11674; MUID:90306003  
 A:Accession: S11674  
 A:Molecule type: DNA  
 A:Residues: 1-421 <KEI>  
 A:Cross-References: EMBL:X54017; NID:g35582; PIDN:CAA37964.1; PID:q1216165  
 A:Note: The authors translated the codon AGG for residue 64 as Thr and CTG for residu  
 R:Vazquez-Levin, M.H.; Revantos, J.; Gordon, J.W.  
 Eur. J. Biochem. 207, 23-26, 1992  
 A:Title: Molecular cloning, sequencing and restriction mapping of the genomic sequenc  
 A:Reference number: S23499; MUID:92331659  
 A:Accession: S23499  
 A:Molecule type: DNA  
 A:Residues: 1-421 <VA2>  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Cross-References: EMBL:M77378  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992  
 R:Keime, S.  
 submitted to the EMBL Data Library, December 1989  
 A:Reference number: S12063  
 A:Accession: S12063  
 A:Molecule type: DNA  
 A:Residues: 1-225, 'R', 227-421 <KEI2>  
 A:Cross-References: EMBL:X54017  
 R:Adham, I.M.; Klemm, U.; Maier, W.M.; Engel, W.  
 Hum. Genet. 84, 125-128, 1990  
 A:Title: Molecular cloning of human preproacrosin cDNA.  
 A:Reference number: A61022; MUID:90128988  
 A:Accession: A61022  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-63, 'T', 65-225, 'V', 227-267, 'R', 269-421 <ADH>  
 R:Baba, T.; Watanabe, K.; Kashiwabara, S.I.; Arai, Y.  
 FEBS Lett. 244, 296-300, 1989  
 A:Title: Primary structure of human proacrosin deduced from its cDNA sequence.  
 A:Reference number: S03330; MUID:89153568  
 A:Accession: S03330  
 A:Molecule type: mRNA  
 A:Residues: 1-63, 'T', 65-119, 'V', 121-165, 'L', 167-267, 'R', 269-344, 'R', 346-421 <BAB>  
 A:Cross-References: EMBL:Y00970; NID:g28325; PIDN:CAA68784.1; PID:g28326  
 C:Genetics:  
 A:Gene: GDB:ACR  
 A:Cross-References: GDB:119645; OMIM:102480  
 A:Map position: 22q13-22qter  
 A:Introns: 26/2; 94/2; 189/2; 237/3  
 C:Superfamily: acrosin; trypsin homology  
 C:Keywords: glycoprotein; hydrolase; serine proteinase; sperm  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-421/Product: acrosin #status predicted <MAY>  
 F:20-42/Product: acrosin light chain #status predicted <LCH>  
 F:43-421/Product: acrosin heavy chain #status predicted <HCH>  
 F:43-285/Domain: trypsin homology <TRY>  
 F:302-379/Region: proline-rich  
 F:22, 210/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:25-154/Disulfide bonds: #status predicted  
 F:29-162/Disulfide bonds: #status predicted  
 F:73-89/Disulfide bonds: #status predicted

F:88,142,240/Active site: His, Asp, Ser #status predicted  
 F:177-246/Disulfide bonds: #status predicted  
 F:209-225/Disulfide bonds: #status predicted  
 F:236-266/Disulfide bonds: #status predicted

Query Match 0.5% Score 7; DB 1; Length 421;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 994 SFARKRLQ 1000  
 |||||  
 DB 383 SFARKRLQ 389

## RESULT 158

A84237  
 hypothetical protein Vng0798h [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: A84237  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A.  
 A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483  
 A:Accession: A84237  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-422 <SNO>  
 A:Cross-references: GB:AE004437; NID:gl0580371; PIDN:AAG19261.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNG0798H

Query Match 0.5% Score 7; DB 2; Length 422;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 DGPFAGG 393  
 |||||  
 DB 251 DGPFAGG 257

## RESULT 159

C83025  
 probable phosphoserine phosphatase PA4960 [imported] - Pseudomonas aeruginosa (strain PA  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: C83025  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: C83025  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-429 <SNO>  
 A:Cross-references: GB:AE004909; GB:AE004091; NID:g9951241; PIDN:AAG08345.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA4960

Query Match 0.5% Score 7; DB 2; Length 429;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 763 FKRLAL 769

Db 263 FKRLAL 269  
 |||||

## RESULT 160

S19722  
 dihydrolipoamide S-acetyltransferase (EC 2.3.1.12) chain E2 - Staphylococcus aureus  
 C:Species: Staphylococcus aureus  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 05-May-2000  
 C:Accession: S19722  
 R:Hemilae, H.  
 Biochim. Biophys. Acta 1129, 119-123, 1991  
 A:Title: Lipoamide dehydrogenase of Staphylococcus aureus: nucleotide sequence and  
 A:Reference number: S19722; MUID:92096451  
 A:Accession: S19722  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-430 <HEM>  
 A:Cross-references: EMBL:X58434; NID:948871; PIDN:CAA41339.1; PID:g581570  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991  
 C:Genetics:  
 A:Start codon: GTG  
 C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology  
 C:Keywords: acyltransferase; coenzyme A  
 F:4-77/Domain: lipoyl/biotin-binding homology <LPB>  
 F:401,405/Active site: His, Asp #status predicted

Query Match 0.5% Score 7; DB 2; Length 430;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 VDAYING 1061  
 |||||  
 DB 158 VDAYING 164

## RESULT 161

T46317  
 hypothetical protein DKFZp434A0612.1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 02-Sep-2000  
 C:Accession: T46317  
 R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, January 2000  
 A:Reference number: Z23035  
 A:Accession: T46317  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-430 <AAA>  
 A:Cross-references: EMBL:AL137749  
 A:Experimental source: adult testis; clone DKFZp434A0612  
 C:Genetics:  
 A:Note: DKFZp434A0612.1  
 C:Superfamily: yeast myo-inositol-1-phosphate synthase

Query Match 0.5% Score 7; DB 2; Length 430;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 |||||  
 DB 202 VSYNHLG 208

## RESULT 162

S47538  
 acrosin (EC 3.4.21.10) precursor - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 26-Dec-1994 #sequence\_revision 10-Nov-1995 #text\_change 22-Jun-1999  
 C:Accession: S47538  
 R:Richardson, R.T.; O'Rand, M.G.



Biochim. Biophys. Acta 1219, 215-218, 1994  
 A:Title: Cloning and sequencing of cDNAs for rabbit preproacrosin and a novel preproacrosin  
 A:Reference number: S47538; MUID:94368861  
 A:Accession: S47538

A:Molecule type: mRNA  
 A:Residues: 1-431 <RIC>  
 A:Cross-References: EMBL:U05204; NID:g451841; PIDN:AAA61630.1; PID:g451842  
 C:Superfamily: acrosin; trypsin homology  
 C:Keywords: glycoprotein; hydrolase; serine proteinase  
 F:40-283/Domain: trypsin homology <TRY>

Query Match 0.5%; Score 7; DB 2; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 994 SPAKRLQ 1000  
 |||||  
 Db 388 SPAKRLQ 394

RESULT 163

E70586  
 hypothetical protein RV2366c - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: E70586  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987

A:Accession: E70586  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA

A:Residues: 1-435 <COL>  
 A:Cross-References: GB:295208; GB:AL123456; NID:g3261747; PIDN:CAB08473.1; PID:g2078033  
 A:Experimental source: strain H37RV  
 C:Genetics:

A:Gene: RV2366c  
 C:Superfamily: hypothetical protein HI0107

Query Match 0.5%; Score 7; DB 2; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 889 IESVFEL 895  
 |||||  
 Db 189 IESVFEL 195

RESULT 164

T16477  
 hypothetical protein F56D2.5 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16477  
 R:Du, Z.

submitted to the EMBL Data Library, August 1994  
 A:Description: The sequence of C. elegans cosmid F56D2.  
 A:Reference number: Z18519  
 A:Accession: T16477

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-437 <DUZ>

A:Cross-References: EMBL:U13644; NID:g532100; PID:g1945502; PIDN:AAB52683.1; GSPDB:GN000  
 A:Experimental source: strain Bristol N2; clone F56D2  
 C:Genetics:  
 A:Gene: CESP:F56D2.5  
 A:Map position: 3

A:Introns: 41/3; 134/3; 185/3; 223/3; 254/2; 287/2; 414/3

Query Match 0.5%; Score 7; DB 2; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 SDKNAEI 236  
 |||||  
 Db 30 SDKNAEI 36

RESULT 165

B83295  
 hypothetical protein PA2794 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: B83295  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
 A:Reference number: A82950; MUID:20437337

A:Accession: B83295  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-438 <STO>  
 A:Cross-References: GB:AE004707; GB:AE004091; NID:g9948876; PIDN:AAG06182.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2794

Query Match 0.5%; Score 7; DB 2; Length 438;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 RVGSGAG 212  
 |||||  
 Db 198 RVGSGAG 204

RESULT 166

T18500  
 hypothetical protein C0765c - malaria parasite (Plasmodium falciparum)  
 C:Species: plasmodium falciparum  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
 C:Accession: T18500  
 R:Lawson, D.; Bowman, S.; Barrell, B.  
 submitted to the EMBL Data Library, August 1997  
 A:Reference number: Z18935

A:Accession: T18500  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA

A:Residues: 1-441 <LAW>  
 A:Cross-References: EMBL:Z98551; NID:el331903; PID:el331921; PIDN:CAB11139.1  
 C:Genetics:  
 A:Map position: 3  
 A:Note: C0765c

Query Match 0.5%; Score 7; DB 2; Length 441;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 827 KISVYVL 833  
 |||||  
 Db 299 KISVYVL 305

RESULT 167  
 A70398

cell division protein ftsW - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C>Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
 C:Accession: A70398  
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oviatt, J.  
 Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A:Reference number: A70300; MUID:98196666  
 A:Accession: A70398  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-448 <KQF>  
 A:Cross-references: GB:AE000725; NID:g2983598; PIDN:AAC07172.1; PID:g2983602; GB:AE00065  
 A:Experimental source: strain VF5  
 C:Genetics:  
 A:Gene: ftsW

Query Match 0.5%; Score 7; DB 2; Length 448;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 444 SLIVENL 450  
 Db 345 SLIVENL 351  
 |||||

RESULT 168  
 S67819  
 GumC protein - Xanthomonas campestris  
 C:Species: Xanthomonas campestris  
 C>Date: 24-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999  
 C:Accession: S67819  
 R:Pollock, T.J.  
 submitted to the EMBL Data Library, March 1995  
 A:Reference number: S67819  
 A:Accession: S67819  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-449 <POL>  
 A:Cross-references: EMBL:U22511; NID:g1172090; PIDN:AAA86371.1; PID:g1172092  
 C:Genetics:  
 A:Gene: gumC

Query Match 0.5%; Score 7; DB 2; Length 449;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 ILNSRLV 982  
 Db 143 ILNSRLV 149  
 |||||

RESULT 169  
 E71044  
 Hypothetical protein PH1645 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
 C:Accession: E71044  
 R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekita, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, N.; Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon Pyrococcus horikoshii strain D-2.  
 A:Reference number: A71000; MUID:98344137  
 A:Accession: E71044  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-450 <KAW>  
 A:Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30757.1; PID:g3258074  
 A:Experimental source: strain OT3  
 A>Note: this accession replaces an interim accession for a sequence replaced by GenBank

## C:Genetics:

A:Gene: PH1645

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1645

Query Match 0.5%; Score 7; DB 2; Length 450;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 QAGIIAN 305

Db 97 QAGIIAN 103

## RESULT 170

D75170

hypothetical protein PAB2013 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C&gt;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: D75170

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome

A:Reference number: A75001

A:Accession: D75170

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-451 &lt;KAW&gt;

A:Cross-references: GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB49443.1; PID:g5

A:Experimental source: strain Orsay

## C:Genetics:

A:Gene: PAB2013

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1645

## Query Match

0.5%; Score 7; DB 2; Length 451;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 QAGIIAN 305

Db 98 QAGIIAN 104

|||||

## RESULT 171

E69056

Conserved hypothetical protein MTH1422 - Methanobacterium thermoautotrophicum (strain

C:Species: Methanobacterium thermoautotrophicum

C&gt;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Nov-1999

C:Accession: E69056

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge,

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jilwan,

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: 1

A:Reference number: A69000; MUID:98037514

A:Accession: E69056

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-457 &lt;MTH&gt;

A:Cross-references: GB:AE000904; GB:AE000666; NID:g2622528; PIDN:AAB85899.1; PID:g26

A:Experimental source: strain Delta H

## C:Genetics:

A:Gene: MTH1422

C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0977

## Query Match

0.5%; Score 7; DB 2; Length 457;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGL 23

|||||

Db 163 LALVGAL 169

RESULT 172

T17480

endo-xylanase homolog PCZA361.14 - Amycolatopsis orientalis

C:Species: Amycolatopsis orientalis

C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000

C:Accession: T17480

R:Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard, N

Chem. Biol. 3, 155-162, 1998

A:Title: Sequencing and analysis of genes involved in the biosynthesis of a vancomycin

A:Reference number: Z18804

A:Accession: T17480

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-462 &lt;VAN&gt;

A:Cross-references: EMBL:AJ223998; NID:el251208; PID:el251219; PIDN:CAAL1771.1

Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 462;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 946 LNTATT 952

|||||

Db 404 LNTATT 410

RESULT 173

S72992

probable phosphorylating protein ureD - Mycobacterium leprae

N:Alternate names: B229\_C2\_192 protein; B229\_C3\_234 protein

C:Species: Mycobacterium leprae

C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 23-Mar-2001

C:Accession: S72992; S72983

R:Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, November 1993

A:Description: Mycobacterium leprae cosmid B229.

A:Reference number: S72988

A:Accession: S72992

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-463 &lt;SMI&gt;

A:Cross-references: EMBL:U00020; NID:9467102; PIDN:AAAL7306.1; PID:9467124

A:Accession: S72983

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 'M', 177-463 &lt;SM2&gt;

A:Cross-references: EMBL:U00020; NID:9467102; PIDN:AAAL7287.1; PID:9467115

C:Genetics:

A:Gene: ureD; ureC

A:Start codon: GTG

C:Superfamily: phosphomannomutase

Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 463;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1051 TSAGVDA 1057

|||||

Db 67 TSAGVDA 73

RESULT 174

A82211

Mut7/nudix family protein VC1342 [imported] - Vibrio cholerae (strain N16961 serogroup C

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: A82211

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: A82211

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-465 &lt;HEI&gt;

A:Cross-references: GB:AE004214; GB:AE003852; NID:99655832; PIDN:AAF94500.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1342

A:Map position: 1

Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 465;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ALVGALV 24

|||||

Db 326 ALVGALV 332

RESULT 175

E82567

GumC protein XF2369 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: E82567

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: E82567

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-467 &lt;SIM&gt;

A:Cross-references: GB:AE004046; GB:AE003849; NID:9107534; PIDN:AAF95168.1; GSPDB:GN

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer

as-Neto, E.; Docena, C.; El-Dorfy, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Tsuchioka, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2369

Query Match

Best Local Similarity 0.5%; Score 7; DB 2; Length 467;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 976 ILNSRLV 982

|||||

Db 159 ILNSRLV 165

RESULT 176

T34504

hypothetical protein ZK1290.12 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T34504  
R:Taich, A.

A:Submitted to the EMBL Data Library, July 1995  
A:Description: The sequence of C. elegans cosmid ZK1290.

A:Reference number: Z21535

A:Accession: T34504

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-484 <TAI>

A:Cross-references: EMBL:U21308; PIDN:AAB93321.1; GSPDB:GN00020; CESP:ZK1290.12

A:Experimental source: strain Bristol N2; clone ZK1290

C:Genetics:

A:Gene: CESP:ZK1290.12

A:Map position: 2

A:Introns: 66/3; 124/2; 181/2; 392/1

Query Match 0.5%; Score 7; DB 2; Length 484;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 VGGFKAS 419

Db 114 VGGFKAS 120

RESULT 177

T32149

hypothetical protein C13A2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jun-2000

C:Accession: T32149

R:Rohlfing, T.; Wohlmann, P.

A:Submitted to the EMBL Data Library, September 1997

A:Description: The sequence of C. elegans cosmid C13A2.

A:Reference number: Z21126

A:Accession: T32149

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-488 <ROH>

A:Cross-references: EMBL:AF022967; PIDN:AAB69873.1; GSPDB:GN00023; CESP:C13A2.1

A:Experimental source: strain Bristol N2; clone C13A2

C:Genetics:

A:Gene: CESP:C13A2.1

A:Map position: 5

A:Introns: 68/2; 115/2; 136/2; 171/3; 209/2; 235/3; 306/3; 351/3

C:Superfamily: Caenorhabditis elegans hypothetical protein F07G11.3

Query Match 0.5%; Score 7; DB 2; Length 488;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 572 HFESEDIG 578

Db 478 HFESEDIG 484

RESULT 178

T28025

hypothetical protein ZK829.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000

C:Accession: T28025

R:Harris, B.

A:Submitted to the EMBL Data Library, May 1996

A:Reference number: Z20458

A:Accession: T28025

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-492 <WII>

A:Cross-references: EMBL:Z73899; PIDN:CAA98077.1; GSPDB:GN00022; CESP:ZK829.9

A:Experimental source: clone ZK829

C:Genetics:

A:Gene: CESP:ZK829.9

A:Map position: 4

A:Introns: 31/1; 148/3; 314/2; 386/2

C:Superfamily: glucose transport protein

Query Match 0.5%; Score 7; DB 2; Length 492;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 918 QTLIDS 924

Db 239 QTLIDS 245

RESULT 179

F96696

protein FLN21.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: F96696

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; A.

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.R.; Creasy, T.H.; Dewar

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kl

A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marz

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: AB6141; MUID:21016719

A:Accession: F96696

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <STO>

A:Cross-references: GB:AE005173; NID:99828628; PIDN:AAG00251.1; GSPDB:GN00141

C:Genetics:

A:Gene: FLN21.12

A:Map position: 1

C:Superfamily: glucose transport protein

Query Match 0.5%; Score 7; DB 2; Length 493;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 594 SLFSGGV 600

Db 107 SLFSGGV 113

RESULT 180

T42758

double-stranded RNA-specific adenosine deaminase homolog - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000

C:Accession: T42758

R:Hough, R.F.; Lingam, A.T.; Bass, B.L.

Nucleic Acids Res 27, 3424-3432, 1999

A:Title: Caenorhabditis elegans mRNAs that encode a protein similar to ADARs derive

A:Reference number: Z22265; MUID:99377169

A:Accession: T42758

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-495 <HOU>

A:Cross-references: EMBL:AF051275; NID:93283022; PIDN:AAC35097.1; PID:93283023

C:Genetics:

A:Map position: 3

Query Match 0.5%; Score 7; DB 2; Length 495;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1158 NALVLKP 1164  
|||||  
Db 225 NALVLKP 231

RESULT 181  
T47715  
hypothetical protein Flii16.170 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 05-May-2000  
C:Accession: T47715  
R:Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: 124473  
A:Accession: T47715  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-497 <BEN>  
A:Cross-references: EMBL:AL161667  
A:Experimental source: cultivar Columbia; BAC clone Flii16  
C:Genetics:  
A:Map position: 3  
A:Introns: 176/1; 390/3; 454/2; 468/2  
A:Note: Flii16.170  
C:Superfamily: Arabidopsis thaliana hypothetical protein Flii16.170

Query Match 0.5%; Score 7; DB 2; Length 497;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 215 ASSTVLT 221  
|||||  
Db 471 ASSTVLT 477

RESULT 182  
S43833  
glyceraldehyde-3-phosphate dehydrogenase (NADP+) (EC 1.2.1.9) - maize  
C:Species: Zea mays (maize)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S43833  
R:Habenicht, A.; Hellman, U.; Cerff, R.  
J. Mol. Biol. 237, 165-171, 1994  
A:Title: Non-phosphorylating GAPDH of higher plants is a member of the aldehyde dehydrogenase  
A:Reference number: S43832; MUID:94180387  
A:Accession: S43833  
A:Molecule type: mRNA  
A:Residues: 1-498 <HAB>  
A:Cross-references: EMBL:X75326; NID:q474407; PIDN:CAA53075.1; PID:q474408  
C:Superfamily: aldehyde dehydrogenase (NADP+); aldehyde dehydrogenase homology  
C:Keywords: NADP; oxidoreductase  
F:56-320/Domain: aldehyde dehydrogenase homology <ALDD>  
F:266,300/Active site: Glu, Cys #status predicted

Query Match 0.5%; Score 7; DB 1; Length 498;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1158 NALVLKP 1164  
|||||  
Db 189 NALVLKP 195

RESULT 183  
S52648  
myo-inositol-1-phosphate synthase (EC 5.5.1.4) - Citrus paradisi  
C:Species: Citrus paradisi  
C:Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 28-Jul-2000

C:Accession: S52648  
R:Holland, D.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: S52648  
A:Accession: S52648  
A:Molecule type: DNA  
A:Residues: 1-507 <HOL>  
A:Cross-references: GB:232632; NID:g602564; PIDN:CAA83565.1; PID:g602565  
C:Genetics:  
A:Gene: INO1  
C:Superfamily: yeast myo-inositol-1-phosphate synthase  
C:Keywords: intramolecular lyase; isomerase

Query Match 0.5%; Score 7; DB 2; Length 507;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1168 VSYNHLG 1174  
|||||  
Db 333 VSYNHLG 339

RESULT 184  
T36370  
probable sensory histidine kinase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T36370  
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, April 1999  
A:Reference number: 221573  
A:Accession: T36370  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-507 <OLI>  
A:Cross-references: EMBL:AL049628; PIDN:CAB40859.1; GSPDB:GN00070; SCOEDB:SCE94.10  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCE94.10

Query Match 0.5%; Score 7; DB 2; Length 507;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ALVGALV 24  
|||||  
Db 65 ALVGALV 71

RESULT 185  
T08436  
myo-inositol-1-phosphate synthase (EC 5.5.1.4) [similarity] - rape  
C:Species: Brassica napus (rape)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 28-Jul-2000  
C:Accession: T08436  
R:Hussain, A.; Bourgeois, J.; Polvi, S.; Tsang, E.; Keller, W.A.; Georges, F.  
submitted to the EMBL Data Library, August 1996  
A:Reference number: 216418  
A:Accession: T08436  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-509 <HUS>  
A:Cross-references: EMBL:U66307; NID:g1513227; PID:g1513228  
C:Function:  
A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-  
C:Superfamily: yeast myo-inositol-1-phosphate synthase  
C:Keywords: intramolecular lyase; isomerase

Query Match 0.5%; Score 7; DB 2; Length 509;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 |||||  
 Db 332 VSYNHLG 338

RESULT 186  
 T04399  
 myo-inositol-1-phosphate synthase (EC 5.5.1.4) - barley  
 C:Species: Hordeum vulgare (barley)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 28-Jul-2000  
 C:Accession: T04399  
 R:Larson, S.R.; Raboy, V.  
 submitted to the EMBL Data Library, March 1998  
 A:Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.  
 A:Reference number: Z14366  
 A:Accession: T04399  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-510 <LAR>  
 A:Cross-references: EMBL:AF056325; NID:g3152730; PIDN:AAC17133.1; PID:g3152731  
 A:Experimental source: cv. Harrington  
 C:Genetics:  
 A:Gene: INO1  
 A:Map position: 4  
 C:Function:  
 A:Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inositol  
 A:Pathway: inositol biosynthesis  
 A>Note: first step  
 C:Superfamily: yeast myo-inositol-1-phosphate synthase  
 C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 0.5%; Score 7; DB 2; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 |||||  
 Db 333 VSYNHLG 339

RESULT 187  
 T01647  
 myo-inositol-1-phosphate synthase (EC 5.5.1.4) - maize  
 C:Species: Zea mays (maize)  
 C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 28-Jul-2000  
 C:Accession: T01647  
 R:Larson, S.R.; Raboy, V.  
 submitted to the EMBL Data Library, March 1998  
 A:Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.  
 A:Reference number: Z14366  
 A:Accession: T01647  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-510 <LAR>  
 A:Cross-references: EMBL:AF056326; NID:g3108052; PIDN:AAC15756.1; PID:g3108053  
 A:Experimental source: strain Early ACR; leaf  
 C:Genetics:  
 A:Gene: INO1  
 C:Function:  
 A:Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inositol  
 A:Pathway: inositol biosynthesis  
 A>Note: NAD cofactor  
 C:Superfamily: yeast myo-inositol-1-phosphate synthase  
 C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 0.5%; Score 7; DB 2; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 |||||  
 Db 333 VSYNHLG 339

RESULT 188  
 S60302  
 myo-inositol-1-phosphate synthase (EC 5.5.1.4) - Spirodela polyrrhiza  
 C:Species: Spirodela polyrrhiza  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 28-Jul-2000  
 C:Accession: S60302  
 R:Smart, C.C.; Fleming, A.J.  
 Plant J. 4, 279-283, 1993  
 A:Title: A plant gene with homology to D-myo-inositol-3-phosphate synthase is rapid  
 A:Reference number: S60302; MUID:94035182  
 A:Accession: S60302  
 A:Molecule type: mRNA  
 A:Residues: 1-510 <SMA>  
 A:Cross-references: EMBL:Z11693; NID:g396209; PIDN:CAA77751.1; PID:g558648  
 C:Genetics:  
 A:Gene: tur1  
 C:Superfamily: yeast myo-inositol-1-phosphate synthase  
 C:Keywords: intramolecular lyase; isomerase

Query Match 0.5%; Score 7; DB 2; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 |||||  
 Db 333 VSYNHLG 339

RESULT 189  
 D84610  
 probable myo-inositol 1-phosphate synthase [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: D84610  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventre  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487  
 A:Accession: D84610  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-510 <STO>  
 A:Cross-references: GB:AE002093; NID:g4567202; PIDN:AAD23618.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g22240  
 A:Map position: 2  
 C:Superfamily: yeast myo-inositol-1-phosphate synthase

Query Match 0.5%; Score 7; DB 2; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 |||||  
 Db 333 VSYNHLG 339

RESULT 190  
 T50021  
 myo-inositol-1-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis  
 N:Alternate names: protein T31P16.160  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 28-Jul-2000  
 C:Accession: T50021

Query Match 0.5%; Score 7; DB 2; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 |||||  
 Db 333 VSYNHLG 339

RESULT 191  
 T50021  
 myo-inositol-1-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis  
 N:Alternate names: protein T31P16.160  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 28-Jul-2000  
 C:Accession: T50021

R:Bevan, M.; Zimmermann, W.; Gruenelsen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; Smith  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25027

A:Accession: T50021  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-510 <BEV>  
A:Cross-references: EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.160  
A:Experimental source: cultivar Columbia; BAC clone T31P16  
C:Genetics:  
A:Gene: ARSP:T31P16.160  
A:Map position: 5  
A:Introns: 63/2; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3  
C:Superfamily: yeast myo-inositol-1-phosphate synthase  
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 0.5%; Score 7; DB 2; Length 510;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
|||||||  
DB 333 VSYNHLG 339

RESULT 191  
I39930  
C:Species: Bacillus thuringiensis  
C:Species: Bacillus thuringiensis  
C:Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 15-Oct-1999  
C:Accession: I39930  
R:Baum, J.A.; Gilbert, M.P.  
J. Bacteriol. 173, 5280-5289, 1991  
A:Title: Characterization and comparative sequence analysis of replication origins from  
A:Reference number: I39930; MUID:91358302  
A:Accession: I39930  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-510 <RES>  
A:Cross-references: GB:M60513; NID:g143272; PIDN:AAA22632.1; PID:g143273  
C:Genetics:  
A:Gene: Ori43

Query Match 0.5%; Score 7; DB 2; Length 510;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 558 LIVKTNG 564  
|||||||  
DB 435 LIVKTNG 441

RESULT 192  
T05017  
myo-inositol-1-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thaliana  
N:Alternate names: protein T19P19.190  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 28-Jul-2000  
C:Accession: T05017  
R:Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hohelsel, J.; Mewes, H.W.;  
submitted to the Protein Sequence Database, April 1998  
A:Reference number: Z15394  
A:Accession: T05017  
A:Molecule type: DNA  
A:Residues: 1-511 <BEV>  
A:Cross-references: EMBL:AL022605  
A:Experimental source: cultivar Columbia; BAC clone T19P19  
C:Genetics:  
A:Map position: 4  
A:Introns: 64/2; 87/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3  
A:Note: T19P19.190

C:Superfamily: yeast myo-inositol-1-phosphate synthase  
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 0.5%; Score 7; DB 2; Length 511;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
|||||||  
DB 334 VSYNHLG 340

RESULT 193  
T10964  
myo-inositol-1-phosphate synthase (EC 5.5.1.4) - kidney bean  
N:Alternate names: 1L-myo-inositol 1-phosphate synthase  
C:Species: Phaseolus vulgaris (kidney bean)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000  
C:Accession: T10964  
R:Wang, X.; Johnson, M.D.  
submitted to the EMBL Data Library, October 1995  
A:Reference number: Z17234  
A:Accession: T10964  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-511 <WAN>  
A:Cross-references: EMBL:U38920; NID:g1066282; PID:g1066283  
A:Experimental source: strain Taylor's horticultural; root  
C:Function:  
A:Description: catalyzes reversible conversion of D-glucose 6-phosphate to 1L-myo-ino  
A:Pathway: myo-inositol biosynthesis  
C:Superfamily: yeast myo-inositol-1-phosphate synthase  
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 0.5%; Score 7; DB 2; Length 511;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
|||||||  
DB 334 VSYNHLG 340

RESULT 194  
T12438  
myo-inositol-1-phosphate synthase (EC 5.5.1.4) - common ice plant  
C:Species: Mesembryanthemum crystallinum (common ice plant)  
C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 28-Jul-2000  
C:Accession: T12438  
R:Ichitani, M.; Majumder, A.L.; Bornhouser, A.; Michalowski, C.B.; Jensen, R.G.; Bohn  
Plant J. 9, 537-548, 1996  
A:Title: Coordinate transcriptional induction of myo-inositol metabolism during enviro  
A:Reference number: Z17518; MUID:96208959  
A:Accession: T12438  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-512 <ISH>  
A:Cross-references: EMBL:U32511; NID:g975987; PIDN:AA03687.1; PID:g975988  
C:Superfamily: yeast myo-inositol-1-phosphate synthase  
C:Keywords: intramolecular lyase; isomerase

Query Match 0.5%; Score 7; DB 2; Length 512;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
|||||||  
DB 335 VSYNHLG 341

RESULT 195  
E82214  
galactoside ABC transporter, ATP-binding protein VCI327 [imported] - Vibrio cholerae (st  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: E82214  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.  
I.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA-Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: E82214  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-514 <HE>  
A:Cross-references: GB:AE004212; GB:AE003852; NID:g9655808; PIDN:AF94485.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCI327  
A:Map position: 1  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 0.5%; Score 7; DB 2; Length 514;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 691 KVATLSV 697  
|||||||  
DB 154 KVATLSV 160

RESULT 196  
A69759  
1-pyruvate-5-carboxylate dehydrogenase homolog ycgN - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: A69759  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerman, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033  
A:Accession: A69759  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-515 <KUN>  
A:Cross-references: GB:Z99105; GB:AL009126; NID:g2632457; PIDN:CAB12115.1; PID:g2632607  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: ycgN  
C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 0.5%; Score 7; DB 2; Length 515;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GGKDTVV 398  
|||||||  
DB 288 GGKDTVV 294

RESULT 197  
A83992  
1-pyruvate-5-carboxylate dehydrogenase BH2737 [imported] - Bacillus halodurans (st  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: A83992  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.;  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans  
A:Reference number: A83650; MUID:20263314  
A:Accession: A83992  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-515 <STO>  
A:Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA006456.1; GSPDB:  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH2737  
C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 0.5%; Score 7; DB 2; Length 515;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GGKDTVV 398  
|||||||  
DB 288 GGKDTVV 294

RESULT 198  
D84142  
1-pyruvate-5-carboxylate dehydrogenase BH3940 [imported] - Bacillus halodurans (st  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: D84142  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.;  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans  
A:Reference number: A83650; MUID:20263314  
A:Accession: D84142  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-515 <STO>  
A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BA007659.1; GSPDB:  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH3940  
C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 0.5%; Score 7; DB 2; Length 515;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GGKDTVV 398  
|||||||  
DB 288 GGKDTVV 294

RESULT 199  
G86335  
hypothetical protein AAF79894.1 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: G86335  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim



C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: G86335  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-519 <STO>  
 A:Cross-references: GB:AE005172; NID:g8778979; PIDN:AAF79894.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: catechol oxidase A2

Query Match 0.5%; Score 7; DB 2; Length 519;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 NNTPSQS 342  
 |||||  
 Db 10 NNTPSQS 16

## RESULT 200

S38921  
 hypothetical protein 9 - phage phi-C31  
 C:Species: phage phi-C31  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
 C:Accession: S38921  
 R:Hartley, N.M.; Murphy, G.O.; Bruton, C.J.; Chater, K.F.  
 submitted to the EMBL Data Library, November 1993  
 A:Reference number: S38912  
 A:Accession: S38921  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-519 <HAR>  
 A:Cross-references: EMBL:X76288; NID:g432610; PIDN:CAAS3920.1; PID:g432620

Query Match 0.5%; Score 7; DB 2; Length 519;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 910 GAQGRDL 916  
 |||||  
 Db 81 GAQGRDL 87

Search completed: August 29, 2001, 09:34:17  
 Job time: 68 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2001, 09:34:29 ; Search time 14.58 Seconds

(without alignments)  
3044.932 Million cell updates/sec

Title: US-09-360-934A-3

Perfect score: 1296

Sequence: 1 MEIQTHRKINRPLVSLALV.....HNLSNIGHFASNLGMRYSF 1296

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 200 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1296	100.0	1296	1 VAC1_HELPY	Q48247 helicobacte
2	110	8.5	1288	1 VAC1_HELPY	Q48245 helicobacte
3	99	7.6	1287	1 VAC2_HELPY	Q48245 helicobacte
4	87	6.7	1290	1 VAC4_HELPY	P55981 helicobacte
5	87	6.7	1291	1 VAC3_HELPY	Q48258 helicobacte
6	62	4.8	1310	1 VAC3_HELPY	Q48253 helicobacte
7	8	0.6	164	1 ATPX_CYAPA	P48085 cyanophora
8	8	0.6	194	1 VATE_SULSO	Q9uww5 sulfobolus
9	8	0.6	334	1 ARG3_BUCAI	P57156 buchnera ap
10	8	0.6	346	1 ARG3_BACSU	P23715 bacillus su
11	8	0.6	637	1 DXS_NEITMA	Q9jw13 bacillus su
12	8	0.6	637	1 DXS_NEITMA	Q9jw13 neisseria m
13	8	0.6	684	1 TC10_YEAST	Q9jxv7 neisseria m
14	7	0.5	90	1 PTSO_ECOLI	P50273 saccharomyc
15	7	0.5	96	1 YDRG_SALTY	P33996 escherichia
16	7	0.5	121	1 VIB2_AGR6	P40864 salmonella
17	7	0.5	121	1 VIB2_AGR6	P09776 agrobacteri
18	7	0.5	144	1 YLX3_CAEL	P05351 agrobacteri
19	7	0.5	151	1 YF12_FOPV	P46499 caenorhabdi
20	7	0.5	159	1 SMPA_TREHY	P36700 fowlpox vir
21	7	0.5	161	1 PRI_ARATH	Q54313 treponema h
22	7	0.5	162	1 MCRM_METH	P33154 arabisopsi
23	7	0.5	177	1 NUSG_BACSU	Q50485 methanobact
24	7	0.5	206	1 N021_SOYBN	Q06795 bacillus su
25	7	0.5	215	1 RK4_ODOSI	P16313 glycine max
26	7	0.5	229	1 HA23_MOUSE	P49546 odontella s
27	7	0.5	235	1 R34E_ARCFU	P14339 mus musculu
28	7	0.5	236	1 BCL2_MOUSE	Q28366 archaeoglob
29	7	0.5	236	1 BCL2_MOUSE	P10417 mus musculu
30	7	0.5	239	1 BCL2_HUMAN	P16313 glycine max
31	7	0.5	247	1 CHIB_LYCES	P49950 rattus norv
32	7	0.5	254	1 HA2R_HUMAN	P10415 homo sapien
33	7	0.5	255	1 HA21_MOUSE	Q05340 lycopersico
					P01903 homo sapien
					P01904 mus musculu

34	7	0.5	255	1 HA22_MOUSE	P04224 mus musculu
35	7	0.5	259	1 HMX2_CHICK	P28362 gallus gall
36	7	0.5	259	1 HMX2_COTJA	P23410 coturnix co
37	7	0.5	263	1 CTBRB_HUMAN	P17538 homo sapien
38	7	0.5	263	1 TYSY_MYCTU	O33306 mycobacteri
39	7	0.5	283	1 LECA_SARPE	P05047 sarcophaga
40	7	0.5	293	1 ALF_STRPN	O65944 streptococ
41	7	0.5	294	1 ARG3_BACST	Q07906 bacillus st
42	7	0.5	295	1 HX71_XENLA	Q04281 xenopus lae
43	7	0.5	299	1 Y175_HELPY	Q92mq7 helicobacte
44	7	0.5	299	1 Y175_HELPY	P56112 helicobacte
45	7	0.5	302	1 PTB_CLOAB	O05624 clostridium
46	7	0.5	318	1 ACCA_ECOLI	P30867 escherichia
47	7	0.5	324	1 ALDX_PIG	P50578 sus scrofa
48	7	0.5	325	1 ACCA_BACSU	Q34847 bacillus su
49	7	0.5	325	1 YGHQ_ECOLI	O46841 escherichia
50	7	0.5	329	1 RBSR_ECOLI	P25551 escherichia
51	7	0.5	330	1 HEM2_HUMAN	P13716 homo sapien
52	7	0.5	334	1 ARG3_ECOLI	P11446 escherichia
53	7	0.5	334	1 HRCB_MYCCA	P71498 mycoplasma
54	7	0.5	344	1 YEHA_ECOLI	P33340 escherichia
55	7	0.5	351	1 RIR2_TREPA	O83092 treponema p
56	7	0.5	353	1 MRAY_HELPY	O25235 helicobacte
57	7	0.5	367	1 YOV4_CAEL	Q22618 caenorhabdi
58	7	0.5	370	1 YXER_BACSU	P54957 bacillus su
59	7	0.5	376	1 OPPC_MYCPN	P75553 mycoplasma
60	7	0.5	389	1 Y823_MYCTU	O53835 mycobacteri
61	7	0.5	395	1 NUSA_HELPY	Q92ja6 helicobacte
62	7	0.5	395	1 NUSA_HELPY	P55977 helicobacte
63	7	0.5	401	1 CBPB_PIG	P05955 sus scrofa
64	7	0.5	403	1 COAT_BOOLV	P12869 boolarra vi
65	7	0.5	410	1 KYK2_DICDI	P18161 dictyosteli
66	7	0.5	410	1 PRI1_CAEL	P34471 caenorhabdi
67	7	0.5	415	1 ACRO_PIG	P08001 sus scrofa
68	7	0.5	421	1 ACRO_HUMAN	P10323 homo sapien
69	7	0.5	430	1 ODP2_STAAU	Q59821 staphylococ
70	7	0.5	431	1 ACRO_RABIT	P48038 oryctolagus
71	7	0.5	434	1 MOTC_RHIME	Q52963 rhizobium m
72	7	0.5	435	1 YN66_MYCTU	O58832 mycobacteri
73	7	0.5	488	1 PS31_ARATH	Q91nu4 arabidopsi
74	7	0.5	488	1 GAPN_MAIZE	O43272 zea mays (m
75	7	0.5	507	1 INOL_CITPA	P42802 citrus para
76	7	0.5	510	1 INOL_SPIO	P42803 spirodela p
77	7	0.5	511	1 INOL_ARATH	P42801 arabidopsi
78	7	0.5	525	1 SP1_RAFA	Q05308 arabidopsi
79	7	0.5	532	1 YMO0_YEAST	Q04458 saccharomyc
80	7	0.5	536	1 HEBX_MOUSE	P20060 mus musculu
81	7	0.5	564	1 ESTJ_HELPY	P12992 heliothis v
82	7	0.5	571	1 IF2_THETH	P48515 thermus aqu
83	7	0.5	647	1 SKO1_YEAST	Q02100 saccharomyc
84	7	0.5	662	1 Y4ID_RHISN	P55487 rhizobium s
85	7	0.5	678	1 CICL_RABIT	P51804 oryctolagus
86	7	0.5	687	1 CICK_HUMAN	P51800 homo sapien
87	7	0.5	687	1 CICK_RABIT	P51803 oryctolagus
88	7	0.5	687	1 CICK_RAT	Q06393 rattus norv
89	7	0.5	687	1 CICL_HUMAN	P51801 homo sapien
90	7	0.5	716	1 PERE_MOUSE	P49290 mus musculu
91	7	0.5	737	1 YNC2_CAEL	P34535 caenorhabdi
92	7	0.5	770	1 YRN9_CAEL	Q09609 caenorhabdi
93	7	0.5	771	1 RIR1_VACCC	P20503 vaccinia vi
94	7	0.5	771	1 RIR1_VACCV	P12848 vaccinia vi
95	7	0.5	774	1 KEMK_MOUSE	Q05512 mus musculu
96	7	0.5	804	1 RIR1_PLAFG	P50647 plasmodium
97	7	0.5	805	1 E2F_DROME	Q27368 drosophila
98	7	0.5	806	1 RIR1_PLAF4	P50648 plasmodium
99	7	0.5	830	1 DYN1_CAEL	P39055 caenorhabdi
100	7	0.5	841	1 PSPI_YEAST	P39056 saccharomyc
101	7	0.5	848	1 DYN3_RAT	Q08877 rattus norv
102	7	0.5	866	1 DYN2_MOUSE	P39057 mus musculu
103	7	0.5	870	1 DYN2_HUMAN	P50570 homo sapien
104	7	0.5	870	1 DYN2_RAT	P39052 rattus norv
105	7	0.5	883	1 DYN_DROME	P27619 drosophila
106	7	0.5	906	1 NUOG_BUCAI	P57257 buchnera ap



```
FT PROPEP ? 1296 POTENTIAL.
SQ SEQUENCE 1296 AA; 139760 MW; 0D1F3F71AB411447 CRC64;

Query Match 100.0%; Score 1296; DB 1; Length 1296;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIQOTHRINRPLVSLALVGVITPQQSHAAFTTIVIPAIVGGIATGTAFTVSGL 60
DB 1 MEIQOTHRINRPLVSLALVGVITPQQSHAAFTTIVIPAIVGGIATGTAFTVSGL 60
QY 61 LSWGLQAEAEANKTPDKPKVRIQAGKGFNEFPKNEYDLYRSLSSKIDGQWDMGNAR 120
DB 61 LSWGLQAEAEANKTPDKPKVRIQAGKGFNEFPKNEYDLYRSLSSKIDGQWDMGNAR 120
QY 121 HYVKGQONKLEVDMDKDAVGYTTLISGLRNETGGDLVNMOKATLRLQFQNGNSFTSYKD 180
DB 121 HYVKGQONKLEVDMDKDAVGYTTLISGLRNETGGDLVNMOKATLRLQFQNGNSFTSYKD 180
QY 181 SADRTTRVDFNAKNISIDNFEINNRVSGAGRKASSTVLTQASEGITSKNAEISLYD 240
DB 181 SADRTTRVDFNAKNISIDNFEINNRVSGAGRKASSTVLTQASEGITSKNAEISLYD 240
QY 241 GATNLASSSVKLGWVWAGRLQYVGYLAPYSYTSINTSKVTGEVNFHNLTVGDKNAQA 300
DB 241 GATNLASSSVKLGWVWAGRLQYVGYLAPYSYTSINTSKVTGEVNFHNLTVGDKNAQA 300
QY 301 GIANKTNICTDLDLWQAGLIITAPPEGYKDKPNNTPSOGAKNDKESAKNDKQESS 360
DB 301 GIANKTNICTDLDLWQAGLIITAPPEGYKDKPNNTPSOGAKNDKESAKNDKQESS 360
QY 361 QNNSNTQVINPNSAQTEVQTOVIDGPFAGGKDTVYNNIRININADGTRVGGFKASL 420
DB 361 QNNSNTQVINPNSAQTEVQTOVIDGPFAGGKDTVYNNIRININADGTRVGGFKASL 420
QY 421 TTNAHLHIGKGVNLSNQASGRSLIVENLTGNITVDGRLRVNNOVGVALAGSANEFE 480
DB 421 TTNAHLHIGKGVNLSNQASGRSLIVENLTGNITVDGRLRVNNOVGVALAGSANEFE 480
QY 481 KAGTDTNGTATFNNDISLGRFVNLKVDATANKGIDTNGGFTLDFSGVTDKVNINK 540
DB 481 KAGTDTNGTATFNNDISLGRFVNLKVDATANKGIDTNGGFTLDFSGVTDKVNINK 540
QY 541 LITASTNVAVKFNINELIVKTNIGSVGEYTHFSDIGSQSRINTVRLTGTGRSLFSGV 600
DB 541 LITASTNVAVKFNINELIVKTNIGSVGEYTHFSDIGSQSRINTVRLTGTGRSLFSGV 600
QY 601 KFKGGEKLVIDEFYSPWNYFDARNIKVEITNKLAFQPGQSPWGTSKLMEFNLTGQNA 660
DB 601 KFKGGEKLVIDEFYSPWNYFDARNIKVEITNKLAFQPGQSPWGTSKLMEFNLTGQNA 660
QY 661 VMDYSQSNLTIOGDFINNOCTINLYVRGKVATLSVGNAAAMFNNDIDSATGYEKLPI 720
DB 661 VMDYSQSNLTIOGDFINNOCTINLYVRGKVATLSVGNAAAMFNNDIDSATGYEKLPI 720
QY 721 KINSAQDLIKNTEHVLKAKIIGVNGVSTGNGISNVNLEQFKERLALYNNNRMDTCV 780
DB 721 KINSAQDLIKNTEHVLKAKIIGVNGVSTGNGISNVNLEQFKERLALYNNNRMDTCV 780
QY 781 VRNTDDIKACMAIGDQSWVNNPDNYKYLICKAMKNIGISKTAGSKISVYILGNSTPTE 840
DB 781 VRNTDDIKACMAIGDQSWVNNPDNYKYLICKAMKNIGISKTAGSKISVYILGNSTPTE 840
QY 841 NCGNTNLTPTNTTSNARSANALQAPPAQPSATPNLVAIQHDFGHTIESVFELANRSK 900
DB 841 NCGNTNLTPTNTTSNARSANALQAPPAQPSATPNLVAIQHDFGHTIESVFELANRSK 900
QY 901 DIDFLYANSQAQGRDLQTLIDSDHAGYARKMIDATSAEITKOLNTATTTLNIAISLE 960
DB 901 DIDFLYANSQAQGRDLQTLIDSDHAGYARKMIDATSAEITKOLNTATTTLNIAISLE 960
QY 961 HKTSGLOTLSLSNAMLNSRLVNLNRRHTNHDFAFKRLQALQKQKFALESAAAEVLIQF 1020
DB 961 HKTSGLOTLSLSNAMLNSRLVNLNRRHTNHDFAFKRLQALQKQKFALESAAAEVLIQF 1020

QY 1021 APKYEKPTNVWANAIGGTSILNNGSNASLYXTSAGVDAYLNGQVEAIVGGFGSYGSYFNN 1080
DB 1021 APKYEKPTNVWANAIGGTSILNNGSNASLYXTSAGVDAYLNGQVEAIVGGFGSYGSYFNN 1080
QY 1081 RANSLNSGANNTNFGVYSRIFANQHEFEDEAQAQALGSDQSSLNFKSALLQDLNQSYYHLA 1140
DB 1081 RANSLNSGANNTNFGVYSRIFANQHEFEDEAQAQALGSDQSSLNFKSALLQDLNQSYYHLA 1140
QY 1141 YSAARTASGYDFAFRNALVLPKPSGVSYNHLGSTNFKSNSTNOVALKNGSSQHLFNA 1200
DB 1141 YSAARTASGYDFAFRNALVLPKPSGVSYNHLGSTNFKSNSTNOVALKNGSSQHLFNA 1200
QY 1201 SANVEARYYYGDTSYFYMNAGVLQEFARHVGSNNAASLNTFFKVNAARNPLNTHARVMMGGE 1260
DB 1201 SANVEARYYYGDTSYFYMNAGVLQEFARHVGSNNAASLNTFFKVNAARNPLNTHARVMMGGE 1260
QY 1261 LKLAKEVFLNLGVVYLHNLISNIGHFASNLGMRYSF 1296
DB 1261 LKLAKEVFLNLGVVYLHNLISNIGHFASNLGMRYSF 1296

RESULT 2
VACA_HELPJ
ID VACA_HELPJ STANDARD; PRT; 1288 AA.
AC Q9ZKW5;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA OR JHP0819.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Gullid B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -!- FUNCTION: INDUCES VACUOLATION OF EUKARYOTIC CELLS. CAUSES
CC ULCERATION AND GASTRIC LESIONS.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001511; AAD06400.1;
KW Cytotoxin; Toxin; Signal.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 ? VACUOLATING CYTOTOXIN.
FT PROPEP 34 1288 POTENTIAL.
SQ SEQUENCE 1288 AA; 139131 MW; 244B159DFC5F32B9 CRC64;

Query Match 8.5%; Score 110; DB 1; Length 1288;
Best Local Similarity 100.0%; Pred. No. 7.3e-105;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 DTKNGTATFNNDISLGRFVNLKVDATANKGIDTNGGFTLDFSGVTDKVNINKLITA 544
```

```

Db 478 DTKNGTATFNNDISLGRFVNLKVDHAHTAFKIDTNGGNTLDFSGVTDKVNINKLITA 537
QY 545 STNVAVKFNINELIVKNGISVGEYTHFSDIGSQSRINTVRLTGTTRS 594
Db 538 STNVAVKFNINELIVKNGISVGEYTHFSDIGSQSRINTVRLTGTTRS 587

RESULT 3
VAC2_HELPY
ID VAC2_HELPY STANDARD; PRT; 1287 AA.
AC Q48245;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49503 / 60190;
RX MEDLINE=94193753; PubMed=8144644;
RA Cover T.L., Tumuru M.K., Cao P., Thompson S.A., Blaser M.J.;
RT "Divergence of genetic sequences for the vacuolating cytotoxin among
RT Helicobacter pylori strains.";
RL J. Biol. Chem. 269:10566-10573(1994).
CC -!- FUNCTION: INDUCES VACUOLATION OF EUKARYOTIC CELLS. CAUSES
CC ULCERATION AND GASTRIC LESIONS.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U05676; AAA17657.1; -
DR Cytootoxin; Toxin; Signal.
KW SIGNAL 1 33
FT CHAIN 34 1287
FT PROPEP 34 1287
FT PROPEP 34 1287
SQ SEQUENCE 1287 AA; 139041 MW; 0007370062FCB71F CRC64;

Query Match 7.6%; Score 99; DB 1; Length 1287;
Best Local Similarity 100.0%; Pred. No. 1.7e-93;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 VGNAAAMFNNDIDSATGFGFKPLIKINSADLIKNTHEVLKAKIIGYGVNSTGTNGISN 756
Db 689 VGNAAAMFNNDIDSATGFGFKPLIKINSADLIKNTHEVLKAKIIGYGVNSTGTNGISN 748
QY 757 VNLSEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 795
Db 749 VNLSEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 787

RESULT 4
VAC4_HELPY
ID VAC4_HELPY STANDARD; PRT; 1290 AA.
AC P55981;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA OR HP0887.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

```

```

OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Klenk H.-P., Gill S., Dougherty B.A.,
RA Fleischmann R.D., Ketchum K.A., Zhou L., Kirkness E.F., Peterson S.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Glodek A.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Hickey E.K.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RT Pylori.;
RL Nature 388:539-547(1997).
CC -!- FUNCTION: INDUCES VACUOLATION OF EUKARYOTIC CELLS. CAUSES
CC ULCERATION AND GASTRIC LESIONS.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000598; AAD07935.1; -
DR TIGR; HP0887; -
KW Cytootoxin; Toxin; Signal.
FT SIGNAL 1 33
FT CHAIN 34 1290
FT PROPEP 34 1290
FT PROPEP 34 1290
SQ SEQUENCE 1290 AA; 139312 MW; F48B23513447ALAC CRC64;

Query Match 6.7%; Score 87; DB 1; Length 1290;
Best Local Similarity 100.0%; Pred. No. 4.5e-81;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTGNTITVDPLRVNNOVGYSALAGSSANPEFKAFTDTKNGTATFNNDISLGRFVNLK 506
Db 442 VENLTGNTITVDPLRVNNOVGYSALAGSSANPEFKAFTDTKNGTATFNNDISLGRFVNLK 501
QY 507 VDAHTANFKGIDTNGGNTLDFSGVT 533
Db 502 VDAHTANFKGIDTNGGNTLDFSGVT 528

RESULT 5
VAC4_HELPY
ID VAC4_HELPY STANDARD; PRT; 1291 AA.
AC Q48258;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=185-44;
RX MEDLINE=94335650; PubMed=8057855;
RA Haas R., Schmitt W.;
RT "Genetic analysis of the Helicobacter pylori vacuolating cytotoxin:
RT structural similarities with the IgA protease type of exported
RT protein.";

```

RL Mol. Microbiol. 12:307-319(1994).  
CC -1- FUNCTION: INDUCES VACUOLATION OF EUKARYOTIC CELLS. CAUSES  
CC ULCEATION AND GASTRIC LESIONS.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z26883; CAA81528.1; -  
KW Cytotoxin; Toxin; Signal.  
FT SIGNAL 1 33 POTENTIAL.  
FT CHAIN 34 ? VACUOLATING CYTOTOXIN.  
FT PROPEP ? 1291 POTENTIAL.  
SQ SEQUENCE 1291 AA; 139635 MW; ECA56A61CAE36669 CRC64;  
-----  
Query Match 6.7%; Score 87; DB 1; Length 1291;  
Best Local Similarity 100.0%; Pred. No. 4.5e-81;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
Qy 447 VENLTGNTVDPVRNNQVGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 506  
Db 442 VENLTGNTVDPVRNNQVGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 501  
-----  
Qy 507 VDAHTANFKGIDTNGGNTLDFSGVT 533  
Db 502 VDAHTANFKGIDTNGGNTLDFSGVT 528  
-----  
RESULT 6  
ID VAC3\_HELPY STANDARD; PRT; 1310 AA.  
AC Q48253;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE VACUOLATING CYTOTOXIN PRECURSOR.  
GN VAC3  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TX30A;  
RX MEDLINE=9535366; PubMed=7629077;  
RA Atherton J.C., Cao P., Peek R.M. Jr., Tummuru M.K., Blaser M.J.,  
RA Cover T.L.;  
RT "Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori.  
RT Association of specific vacA types with cytotoxin production and  
RT peptic ulceration."  
RL J. Biol. Chem. 270:17771-17777(1995).  
CC -1- FUNCTION: INDUCES VACUOLATION OF EUKARYOTIC CELLS. CAUSES  
CC ULCEATION AND GASTRIC LESIONS.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U29401; AAA86834.1; -  
KW Cytotoxin; Toxin; Signal.  
FT SIGNAL 1 30 POTENTIAL.  
FT CHAIN 31 ? VACUOLATING CYTOTOXIN.  
-----

FT PROPEP ? 1310 POTENTIAL.  
SQ SEQUENCE 1310 AA; 141988 MW; IBC21FE3D435F981 CRC64;  
-----  
Query Match 4.8%; Score 62; DB 1; Length 1310;  
Best Local Similarity 100.0%; Pred. No. 3.3e-55;  
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
Qy 1082 ANSLSGANTNFGVYSRIFANQHEFDEAQAQSGALGSDQSSLNFKSALLQDLNOSYHYLAY 1141  
Db 1097 ANSLSGANTNFGVYSRIFANQHEFDEAQAQSGALGSDQSSLNFKSALLQDLNOSYHYLAY 1156  
-----  
Qy 1142 SA 1143  
Db 1157 SA 1158  
-----  
RESULT 7  
ID ATPX\_CYAPA STANDARD; PRT; 164 AA.  
AC P48085;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE ATP SYNTHASE B' CHAIN (EC 3.6.1.34) (SUBUNIT II).  
GN ATPG.  
OS Cyanophora paradoxa.  
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.  
OX NCBI\_TaxID=2762;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LB555 / PRINGSHEIM;  
RA Stirewalt V.L., Michalowski C.B., Luffelhardt W., Bohnert H.J.,  
RA Bryant D.A.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBUNIT: THIS IS ONE OF THE FOUR CHAINS OF THE NONENZYMATIC  
CC COMPONENT (CF(0) SUBUNIT) OF THE CHLOROPLAST ATPASE COMPLEX.  
CC -1- SUBCELLULAR LOCATION: CYANELLE THYLAKOID MEMBRANE.  
CC -1- SIMILARITY: THE B'-SUBUNIT IS A DIVERGED AND DUPLICATED FORM OF  
CC B FOUND IN PLANTS AND PHOTOSYNTHETIC BACTERIA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U30821; AAA81256.1; -  
DR Mendel; 7860; CYAPA:atpg;1.  
DR InterPro; IPR002146; -  
DR Pfam; PF00430; ATP-synt\_B; 1.  
KW Hydrogen ion transport; Transmembrane; CF(0); Cyanelle.  
SQ SEQUENCE 164 AA; 18568 MW; 49AAACE15AF010D7C CRC64;  
-----  
Query Match 0.6%; Score 8; DB 1; Length 164;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
Qy 715 FYKPLIKI 722  
Db 46 FYKPLIKI 53  
-----  
RESULT 8  
ID VATE\_SULSO STANDARD; PRT; 194 AA.  
AC Q9UWWS;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)

01-OCT-2000 (Rel. 40, Last annotation update)  
DE V-TYPE APP SYNTHASE SUBUNIT E (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT E).  
GN ATPE.  
OS Sulfolobus solfataricus.  
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
OX NCBI\_TaxID=2287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 1617 / P2;  
RX MEDLINE=20165948; PubMed=10701121;  
RA Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,  
Gastonier F., Curtis B., Duguet M., Erasuo G., Faguy D.,  
Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,  
Kushwaha N., Lafleur E., Medina N., Peng X., Penny S.L., She O.,  
St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,  
Ragan M.A., Sensen C.W.;  
RT "Gene content and organization of a 281-kbp contig from the genome of  
the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";  
RL Genome 43:116-136(2000).  
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
CC GRADIENT ACROSS THE MEMBRANE.  
CC -1- SIMILARITY: BELONGS TO THE V-ATPASE E SUBUNIT FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: Y18930; CAB57738.1;  
KW Hydrolyase; ATP synthesis; Hydrogen ion transport.  
SQ SEQUENCE 194 AA; 22633 MW; 057157C921ED8449 CRC64;  
  
Query Match 0.6%; Score 8; DB 1; Length 194;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 555 INELVKT 562  
Db 48 INELVKT 55  
|||||||  
  
RESULT 9  
ARGC\_BUCAI  
ID ARGC\_BUCAI STANDARD; PRT; 334 AA.  
AC P57156;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (EC 1.2.1.38) (AGPR) (N-  
DE ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE).  
GN ARGC OR BU048.  
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
OS symbiotic bacterium).  
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
OX NCBI\_TaxID=118099;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TOKYO 1998;  
RX MEDLINE=20445173; PubMed=10993077;  
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
RT "Genome sequence of the endocellular bacterial symbiont of aphids  
RT Buchnera sp. APS.";  
RL Nature 407:81-86(2000).  
CC -1- CATALYTIC ACTIVITY: N-ACETYL-L-GLUTAMATE 5-SEMIALDEHYDE + NADP(+) +  
CC + ORTHOPHOSPHATE = N-ACETYL-5-GLUTAMYL PHOSPHATE + NADPH.  
CC -1- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.  
CC -1- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: AP001118; BAB12771.1;  
DR PROSITE; PS01224; ARGC: 1  
KW Arginine biosynthesis; Oxidoreductase; NADP.  
FT ACT\_SITE 154 154 BY SIMILARITY.  
SQ SEQUENCE 334 AA; 37815 MW; 1032E60048AFA90A CRC64;  
  
Query Match 0.6%; Score 8; DB 1; Length 334;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 209 SGAGRKAS 216  
Db 187 SGAGRKAS 194  
|||||||  
  
RESULT 10  
ARGC\_BACSU  
ID ARGC\_BACSU STANDARD; PRT; 346 AA.  
AC P23715; P70953; O08146;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (EC 1.2.1.38) (AGPR) (N-  
DE ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE).  
GN ARGC.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / EMG50;  
RX MEDLINE=90356403; PubMed=2117746;  
RA Smith M.C.M., Mountain A., Baumberg S.;  
RT "Nucleotide sequence of the Bacillus subtilis argC gene encoding N-  
RT acetylglutamate-gamma-semialdehyde dehydrogenase.";  
RL Nucleic Acids Res. 18:4595-4595(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=94297722; PubMed=8025667;  
RA O'Reilly M., Devine K.M.;  
RT "Sequence and analysis of the citrulline biosynthetic operon argC-F  
RT from Bacillus subtilis.";  
RL Microbiology 140:1023-1025(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=97177785; PubMed=9025291;  
RA Levine A., Vannier F., Roche B., Autret S., Mavel D., Seror S.J.;  
RT "A 10.3 kbp segment from nprB to argC at the 102 degrees region of  
RT the Bacillus subtilis chromosome.";  
RL Microbiology 143:175-177(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98015415; PubMed=9353931;  
RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;  
RT "Sequencing of regions downstream of addA (98 degrees) and citG (289  
RT degrees) in Bacillus subtilis.";  
RL Microbiology 143:3305-3308(1997).  
RN [5]  
RP SEQUENCE OF 1-56 FROM N.A.  
RX MEDLINE=87192000; PubMed=3106155;  
RA Smith M.C.M., Mountain A., Baumberg S.;

```

RT *sequence analysis of the Bacillus subtilis argC promoter region.;
RL Gene 49:53-60(1986).
CC -1- CATALYTIC ACTIVITY: N-ACETYL-L-GLUTAMATE 5-SEMIALDEHYDE + NADP(+)
CC + ORTHOPHOSPHATE -> N-ACETYL-5-GLUTAMYL PHOSPHATE + NADPH.
CC -1- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52834; CAA37016.1; -
DR EMBL; Z26919; CAA81543.1; -
DR EMBL; Z79580; CAB01842.1; -
DR EMBL; Y09476; CAA70638.1; -
DR EMBL; Z99110; CAB12976.1; -
DR EMBL; Z99109; CAB12960.1; -
DR EMBL; M15420; AAA22248.1; -
DR PIR; S12592; S12592.
DR Subtilist; BG10191; argC.
DR InterPro; IPR000534; -
DR InterPro; IPR000706; -
DR Pfam; PF0118; Semialdehyde_dh; 1.
DR PROSITE; PS01224; ARGCG; 1.
KW Arginine biosynthesis; Oxidoreductase; NADP.
FT ACT_SITE 149 149
FT CONFLICT 235 235 F -> V (IN REF. 3 AND 4).
FT CONFLICT 341 341 MISSING (IN REF. 1).
SQ SEQUENCE 346 AA; 38121 MW; 9E22F2AB31B7542B CRC64;

Query Match 0.6%; Score 8; DB 1; Length 346;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKAS 216
DB 181 SGAGRKAS 188

RESULT 11
DXX_NEIMA
ID DXX_NEIMA STANDARD; PRT; 637 AA.
AC Q9JWL3;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 1-DEOXY-D-XYLULOSE 5-PHOSPHATE SYNTHASE (EC 2.2.-.-) (1-DEOXYXYLULOSE-
DE 5-PHOSPHATE SYNTHASE) (DXP SYNTHASE) (DXPS).
GN DXS OR NWA0589.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT *Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.
RL Nature 404:502-506(2000).
CC -1- FUNCTION: CATALYZES THE ACYLON CONDENSATION REACTION BETWEEN C
CC ATOMS 2 AND 3 OF PYRUVATE AND GLYCERALDEHYDE 3-PHOSPHATE TO YIELD
CC 1-DEOXY-D-XYLULOSE-5-PHOSPHATE (DXP) (BY SIMILARITY).

```

```

CC -1- COFACTOR: THIAMINE PYROPHOSPHATE (BY SIMILARITY).
CC -1- PATHWAY: DEOXYXYLULOSE-5-PHOSPHATE PATHWAY (DXP) OF ISOPRENOID
CC BIOSYNTHESIS; FIRST STEP. BIOSYNTHETIC PATHWAY TO THIAMINE AND
CC PYRIDOXOL; FIRST STEP.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY. DXS SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL162753; CAB83880.1; -
DR InterPro; IPR000360; -
DR PROSITE; PS00801; TRANSKETOLASE.1; 1.
DR PROSITE; PS00802; TRANSKETOLASE.2; 1.
KW Transferase; Flavoprotein; Thiamine pyrophosphate;
KW Isoprene biosynthesis; Thiamine biosynthesis.
SQ SEQUENCE 637 AA; 68720 MW; 3B2BBD01AAD182F5 CRC64;

Query Match 0.6%; Score 8; DB 1; Length 637;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 GHFASNLG 1291
DB 44 GHFASNLG 51

RESULT 12
DXX_NEIMB
ID DXX_NEIMB STANDARD; PRT; 637 AA.
AC Q9JXV7;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 1-DEOXY-D-XYLULOSE 5-PHOSPHATE SYNTHASE (EC 2.2.-.-) (1-DEOXYXYLULOSE-
DE 5-PHOSPHATE SYNTHASE) (DXP SYNTHASE) (DXPS).
GN DXS OR NMB1867.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT *Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.
RL Science 287:1809-1815(2000).
CC -1- FUNCTION: CATALYZES THE ACYLON CONDENSATION REACTION BETWEEN C
CC ATOMS 2 AND 3 OF PYRUVATE AND GLYCERALDEHYDE 3-PHOSPHATE TO YIELD
CC 1-DEOXY-D-XYLULOSE-5-PHOSPHATE (DXP) (BY SIMILARITY).
CC -1- COFACTOR: THIAMINE PYROPHOSPHATE (BY SIMILARITY).
CC -1- PATHWAY: DEOXYXYLULOSE-5-PHOSPHATE PATHWAY (DXP) OF ISOPRENOID
CC BIOSYNTHESIS; FIRST STEP. BIOSYNTHETIC PATHWAY TO THIAMINE AND
CC PYRIDOXOL; FIRST STEP.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY. DXS SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AE002536; AAF42201.1; -  
 DR TIGR; NMB1867; -  
 DR InterPro; IPR000360; -  
 DR PROSITE; PS00801; TRANSKETOLASE\_1; 1.  
 DR PROSITE; PS00802; TRANSKETOLASE\_2; 1.  
 KW Transferrase; Flavoprotein; Thiamine pyrophosphate;  
 KW Isoprene biosynthesis; Thiamine biosynthesis.  
 SQ SEQUENCE 637 AA; 68749 MW; DF5PD396CF6FAF51 CRC64;

Query Match 0.6%; Score 8; DB 1; Length 637;

Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 GHEASNLG 1291

Db 44 GHFASNLG 51

RESULT 13

TC10\_YEAST  
 ID TC10\_YEAST STANDARD; PRT; 684 AA.

AC P50273;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE TCM10 PROTEIN.

GN TCM10 OR YDR350C OR D9476.9.

OS Saccharomyces cerevisiae (Baker's Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MH125;

RA Zhang Y., Robinson K.M., Lemire B.D.;

RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,

RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J.,

RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,

RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,

RA Ridsen E., Meneses S., Miller N., Nhan M., Pauley A., Peluso D.,

RA Rifken L., Riles L., Faich A., Trevasakis E., Vignati D.,

RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U32306; AAA74031.1; -

DR EMBL; U28372; AAB64786.1; -

DR SGD; S0002758; TCM10.

CC CONFLICT 592 684

FT GARSWYNKILFGFEIRHMALQIIOKQGFPPKPNFDTLL

FT TELVNNIKPEPDTSTLTDEDMEDGKPRFNDYDKCT

FT NIRETLKSLN -> EHAPGTIRYSSALKSGIWL (IN

FT REF. 2).

FT SQ SEQUENCE 684 AA; 79755 MW; A88992848F8F49A CRC64;

Query Match 0.6%; Score 8; DB 1; Length 684;

Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LYRSLLSS 107

Db 452 LYRSLLSS 459

RESULT 14

PTSO\_ECOLI

ID PTSO\_ECOLI STANDARD; PRT; 90 AA.

AC P33996;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE PHOSPHOCARRIER PROTEIN NPR (NITROGEN RELATED HPR).

GN PTSO OR NPR OR RPOR.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia

OX NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RC STRAIN=K12 / W3110;

RX MEDLINE=95181483; PubMed=7876255;

RA Powell B.S., Court D.L., Inada T., Nakamura Y., Michotey V.,

RA Cui X., Reizer A., Saier M.H. Jr., Reizer J.;

RT "Novel proteins of the phosphotransferase system encoded within the

RT rpoN operon of Escherichia coli. Enzyme IIA<sup>NTR</sup> affects growth on

RT organic nitrogen and the conditional lethality of an *erats* mutant.,"

RL J. Biol. Chem. 270:4822-4839(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=94297724; PubMed=8025669;

RA Jones D.H.A., Franklin C.F.H., Thomas C.M.;

RT "Molecular analysis of the operon which encodes the RNA polymerase

RT sigma factor sigma 54 of Escherichia coli.,"

RL Microbiology 140:1035-1043(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RA Plunkett G. III;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: SEEMS TO HAVE A ROLE IN LINKING CARBON AND NITROGEN

CC ASSIMILATION. PROBABLY ACT IN A REGULATORY CAPACITY AND COULD

CC CONTROL THE STATE OF PHOSPHORYLATION OF IIA-NTR (PTS).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).

CC -1- PTM: PROBABLY PHOSPHORYLATED BY A YET UNCHARACTERIZED ENZYME I.

CC -1- SIMILARITY: TO ALL OTHER HPR OR HPR DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U12684; AAB60167.1; -

DR EMBL; 227094; CAAB1621.1; -

DR EMBL; U18997; AAA58008.1; -

DR EMBL; AE000400; AAC76238.1; -

DR PIR; S38619; S38619.

DR HSSP; P08877; 2HPR.

DR EcoGene; EG12147; ptsO.

DR InterPro; IPR001020; -

DR Pfam; PF00381; PTS-Hpr; 1.

DR PROSITE; PS00369; PTS\_HPR\_HIS; 1.

DR PROSITE; PS00589; PTS\_HPR\_SER; 1.

DR Phosphotransferase system; Phosphorylation.

FT MOD\_RES 16 16

FT MOD\_RES 48 48

FT PHOSPHORYLATION (BY SIMILARITY).

FT PHOSPHORYLATION (BY SIMILARITY).

```
SQ SEQUENCE 90 AA; 9810 MW; 12A8EFEB514F6015 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 629 VEITNKL 635
DB 7 VEITNKL 13

RESULT 15
YDFG_SALTY STANDARD; PRT; 96 AA.
AC P40864;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL OXIDOREDUCTASE IN DCP 3'REGION (EC 1.-.-.-) (FRAGMENT).
GN YDFG.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92165738; PubMed=1537804;
RA Miller C.G., Hamilton S.;
RT "Cloning and nucleotide sequence of the Salmonella typhimurium dcp
gene encoding dipeptidyl carboxypeptidase.";
RL J. Bacteriol. 174:1626-1630(1992).
RN [2]
RP IDENTIFICATION.
RX MEDLINE=95004589; PubMed=7920643;
RA Robison K., Gilbert W., Church G.M.;
RT "Large scale bacterial gene discovery by similarity search.";
RL Nat. Genet. 7:205-214(1994).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
(SDR) FAMILY. STRONG, TO OTHER BACTERIAL HOMOLOGS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84575; ; NOT_ANNOTATED_CDS.
DR HSSP; P14061; 1FDW.
DR StyGene; SG10499; ydfg.
DR InterPro; IPR002198; -.
DR PROSITE; PS00061; ADH_SHORT; PARTIAL.
KW Hypothetical protein; Oxidoreductase.
FT NON_TER 96
SQ SEQUENCE 96 AA; 10437 MW; 275A7E29C27ED1A7 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 998 RLQALKD 1004
DB 36 RLQALKD 42

RESULT 16
VIB2_AGR76
ID VIB2_AGR76 STANDARD; PRT; 121 AA.
AC P09776;
DT 01-MAR-1989 (Rel. 10, Created)

SQ SEQUENCE 90 AA; 9810 MW; 12A8EFEB514F6015 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 LSLSNAM 975
DB 14 LSLSNAM 20

RESULT 17
VIB2_AGR79
ID VIB2_AGR79 STANDARD; PRT; 121 AA.
AC P05351;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE VIB2 PROTEIN PRECURSOR.
GN VIB2.
OS Agrobacterium tumefaciens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=362;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88247765; PubMed=2837739;
RA Thompson D.V., Melchers L.S., Idler K.B., Shilperoort R.A.,
RA Hooykaas P.J.J.;
RT "Analysis of the complete nucleotide sequence of the Agrobacterium
tumefaciens virB operon.";
RL Nucleic Acids Res. 16:4621-4636(1988).
CC -1- FUNCTION: VIRB PROTEINS ARE SUGGESTED TO ACT AT THE BACTERIAL
SURFACE AND THERE PLAY AN IMPORTANT ROLE IN DIRECTING T-DNA
TRANSFER TO PLANT CELLS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: X06826; CAA29973.1; -  
DR PIR: S00778; B2AG55.  
KW Crown gall tumor; Plasmid; Signal.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 121 VIRB2 PROTEIN.  
SQ SEQUENCE 121 AA; 12288 MW; 49A59C8E04AC6913 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 121;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 969 LSLSNAM 975  
Db 14 LSLSNAM 20  
|||||

## RESULT 18

YLX3 CAEEL STANDARD; PRT; 144 AA.  
AC P46499;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE HYPOTHETICAL 16.2 KDA PROTEIN P23F12.3 IN CHROMOSOME III.  
GN F23F12.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Du 2.;  
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: U12965; AAA20605.1; -  
DR WormPep; F23F12.3; CE01250.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 11 31 POTENTIAL.  
FT TRANSMEM 97 117 POTENTIAL.  
FT TRANSMEM 121 141 POTENTIAL.  
SQ SEQUENCE 144 AA; 16217 MW; 014E0CB5328E9B09 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 144;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ALVGALV 24  
Db 107 ALVGALV 113  
|||||

## RESULT 19

YF12\_FOWPV STANDARD; PRT; 151 AA.  
ID YF12\_FOWPV

AC P36700;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE HYPOTHETICAL PROTEIN IN F12 HOMOLOG 3' REGION (ORF3) (FRAGMENT).  
OS Fowlpox virus.  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Avipoxvirus.  
OX NCBI\_TaxID=10261;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93139784; PubMed=8380837;  
RA Ogawa R., Calvert J.G., Yanagida N., Nazerian K.;  
RT "Insertional inactivation of a fowlpox virus homologue of the  
RL vaccinia virus F12L gene inhibits the release of enveloped virions.";  
J. Gen. Virol. 74:55-64(1993).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: M88588; AAA47188.1; -  
DR PIR: S27935; S27935.  
DR PIR: PQ0506; PQ0506.  
KW Hypothetical protein.  
FT NON\_TER 151  
SQ SEQUENCE 151 AA; 16892 MW; A37699544B48FCBF CRC64;

Query Match 0.5%; Score 7; DB 1; Length 151;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 531 GVTDKVN 537  
Db 69 GVTDKVN 75  
|||||

## RESULT 20

SMPA\_TREHY STANDARD; PRT; 159 AA.  
AC Q54313; 1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE 16 KDA OUTER MEMBRANE LIPOPROTEIN PRECURSOR.  
GN SMPA.  
OS Treponema hyodysenteriae (Serpulina hyodysenteriae).  
OC Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.  
OX NCBI\_TaxID=159;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=P18A;  
RX MEDLINE=93162807; PubMed=8432595;  
RA Thomas W., Sellwood R.;  
RT "Molecular cloning, expression, and DNA sequence analysis of the gene  
RT that encodes the 16-kilodalton outer membrane lipoprotein of  
RT Serpulina hyodysenteriae.";  
RL Infect. Immun. 61:1136-1140(1993).  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
CC ANCHOR.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

```

CC EMBL; X68401; CAA48467.1;
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1;
KW Repeat; Outer membrane; Lipoprotein; Signal.
FT SIGNAL 21 PROBABLE.
FT CHAIN 22 159 16 KDA OUTER MEMBRANE LIPOPROTEIN.
FT LIPID 22 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 159 AA; 16838 MW; EF8C976DE09F32AE CRC64;

Query Match 0.5%; Score 7; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 515 KGIDTGN 521
DB 100 KGIDTGN 106

RESULT 21
PRL_ARATH
ID PRL_ARATH STANDARD; PRT; 161 AA.
AC P33154;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PATHOGENESIS-RELATED PROTEIN 1 PRECURSOR (PR-1).
GN AT2G14610 OR T6B13.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP STRAIN=CV. LANDSBERG ERECTA; TISSUE=Leaf;
RX MEDLINE=93005717; PubMed=1392589;
RA Uknes S., Mauch-Mani B., Moyer M., Potter S., Williams S.,
RA Dincher S., Chandler D., Slusarenko A., Ward E., Ryals J.;
RT "Acquired resistance in Arabidopsis.",
RL Plant Cell 4:645-656(1992).
[2]
SEQUENCE FROM N.A.
RP STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Unayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.",
RL Nature 402:761-768(1999).
CC -1- FUNCTION: PARTIALLY RESPONSIBLE FOR ACQUIRED PATHOGEN RESISTANCE.
CC -1- SUBCELLULAR LOCATION: ACCUMULATES IN THE APOPLAST BEFORE
SECRETION.
CC -1- INDUCTION: INDUCED BY 2,6-DICHLOROISOCOTINIC ACID (INA) AND
SALICYLIC ACID (POSSIBLY AN ENDOGENOUS SIGNAL FOR ACQUIRED
RESISTANCE). STRONGLY INDUCED BY PATHOGEN INFECTION.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M90508; AAA32863.1;

```

```

DR EMBL; AC005398; AAC69381.1;
DR PIR; J01693; J01693.
DR HSP; P04284; ICPE.
DR InterPro; IPR001283; -.
DR Pfam; PF00188; SCP; 1.
DR PRINTS; PRO0837; V5TPXLIKE.
DR PROSITE; PS10009; SCP_AG5_PRL_SC7_1; 1.
DR PROSITE; PS10101; SCP_AG5_PRL_SC7_2; 1.
KW Signal; Pathogenesis-related protein.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 161 PATHOGENESIS-RELATED PROTEIN 1.
FT MOD_RES 27 27 PYRROLIDONE CARBOXYLIC ACID
(BY SIMILARITY).
FT DISULFID 70 138 BY SIMILARITY.
FT DISULFID 113 117 BY SIMILARITY.
FT DISULFID 133 147 BY SIMILARITY.
SQ SEQUENCE 161 AA; 17677 MW; 898B0FF6547C3F84 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGALV 24
DB 15 ALVGALV 21

RESULT 22
MCRW_METH
ID MCRW_METH STANDARD; PRT; 162 AA.
AC Q50485; O27203;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE METHYL-COENZYME M REDUCTASE II OPERON PROTEIN D.
GN MRED OR MTH1131.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
[1]
SEQUENCE FROM N.A.
RP STRAIN=DELTA H;
RX MEDLINE=95014084; PubMed=7929010;
RA Pihl T.D., Sharma S., Reeve J.N.;
RT "Growth phase-dependent transcription of the genes that encode the
two methyl coenzyme M reductase isoenzymes and N5-
methyltetrahydromethanopterin:coenzyme M methyltransferase in
Methanobacterium thermoautotrophicum delta H.",
RL J. Bacteriol. 176:6384-6391(1994).
[2]
SEQUENCE FROM N.A.
RP STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
delta: functional analysis and comparative genomics.",
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- FUNCTION: THIS ENZYME COMPLEX CATALYZES THE FINAL STEP IN
METHANOGENESIS, WHICH IS THE TERMINAL STEP OF ANAEROBIC
DEGRADATION OF BIOMASS.
CC -1- SUBUNIT: MCR IS COMPOSED OF THREE SUBUNITS: ALPHA, BETA, AND
GAMMA. THE FUNCTION OF PROTEINS C AND D IS NOT KNOWN.
CC -1- DEVELOPMENTAL STAGE: THERE ARE TWO MCR COMPLEXES IN THIS BACTERIA.
MCR II IS EXPRESSED IN THE EARLY GROWTH PHASE. LATE GROWTH CELLS
CONTAINS MOSTLY MCR I.

```

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: U09990; AAA73437.1; -  
EMBL: AE000883; AAB85620.1; -  
KW Methanogenesis; Multigene family.  
FT CONFLICT 44 52 PSIPDRVYY -> HQYPTVT (IN REF. 1).  
FT CONFLICT 58 59 TE -> OR (IN REF. 1).  
SQ SEQUENCE 162 AA; 18619 MW; 3D90667C7BB1D164 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 162;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 394 KDTVWNI 400  
Db 155 KDTVWNI 161  
|||||

RESULT 23  
NUSG\_BACSU STANDARD; PRT; 177 AA.  
AC Q06795;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE TRANSCRIPTION ANTI-TERMINATION PROTEIN NUSG.  
GN NUSG.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95058172; PubMed=7968510;  
RA Jeong S., Yoshikawa H., Takahashi H.;  
RT Isolation and characterization of the secE homologue gene of  
RT Bacillus subtilis.;  
RL Mol. Microbiol. 10:133-142(1993).  
CC -!- FUNCTION: INFLUENCES TRANSCRIPTION TERMINATION AND  
CC ANTI-TERMINATION. ACTS AS A COMPONENT OF THE TRANSCRIPTION COMPLEX,  
CC AND INTERACTS WITH THE TERMINATION FACTOR RHO AND RNA POLYMERASE  
CC (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE NUSG FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: D13303; BAA02560.1; -  
DR EMBL; Z99104; CAB11877.1; -  
DR PIR; S39859; S39859.  
DR PIR; S40071; S40071.  
DR Subtilist; BG10162; nusG.  
DR InterPro; IPR001062; -  
DR PRINTS; PR00338; NUSGTNSCPFCT.  
DR PROSITE; PS01014; NUSG; 1.  
KW Transcription termination.  
SQ SEQUENCE 177 AA; 20126 MW; C9F2557AC35262EF CRC64;

Query Match 0.5%; Score 7; DB 1; Length 177;

Best Local Similarity 100.0%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 385 VIDGPFA 391  
Db 132 VIDGPFA 138  
|||||

RESULT 24  
NO21\_SOYBN STANDARD; PRT; 206 AA.  
AC P16313;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE NODULIN 21 (N-21).  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
OC Fabales; Fabaceae; Papilionoideae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. PRIZE; TISSUE=Root nodules;  
RX MEDLINE=91346633; PubMed=2102825;  
RA Delauney A.J., Cheon C.-I., Verma D.P.S.;  
RT "A nodule-specific sequence encoding a methionine-rich polypeptide,  
RT nodulin-21.";  
RL Plant Mol. Biol. 14:449-451(1990).  
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SYMBIOTIC NITROGEN  
CC FIXATION.  
CC -!- DEVELOPMENTAL STAGE: ABUNDANT DURING NODULE DEVELOPMENT.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: X16488; CAA34506.1; -  
DR PIR; S08632; S08632.  
KW Nodulation; Nitrogen fixation.  
SQ SEQUENCE 206 AA; 21833 MW; E2468DD2D4332AE5 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 789 ACGMAIG 795  
Db 100 ACGMAIG 106  
|||||

RESULT 25  
RK4\_ODOSI STANDARD; PRT; 215 AA.  
ID RK4\_ODOSI  
AC P49546;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L4.  
GN RPL4.  
OS Odontella sinensis.  
OG Chloroplast.  
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;  
OC Biddulphiophycidae; Eupodiscales; Eupodiscaeae; Odontella.  
OX NCBI\_TaxID=2839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;

RT "The chloroplast genome of a chlorophyll a+c-containing alga,  
 RL *Odontella sinensis*.", 13:336-342(1995).  
 RC Plant Mol. Biol. Rep. 13:336-342(1995).  
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY AND SPECIFICALLY TO 23S  
 CC RNA (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE LAP FAMILY OF RIBOSOMAL PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Z67753; CAA91648.1; .  
 DR Mendel; 9492; ODOsi.rpl4.1.  
 DR InterPro; IPR002136; .  
 DR Pfam; PF00573; Ribosomal\_L4; 1.  
 KW Ribosomal protein; rRNA-binding; Chloroplast.  
 SQ SEQUENCE 215 AA; 24349 MW; B34897A571453559 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 215;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 624 RNKNVE 630  
 DB 175 RNKNVE 181  
 |||||

RESULT 26  
 HA23\_MOUSE  
 ID HA23\_MOUSE STANDARD; PRT; 229 AA.  
 AC P14439;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-NOV-1991 (Rel. 20, Last annotation update)  
 DE H-2 CLASS II HISTOCOMPATIBILITY ANTIGEN, E-U ALPHA CHAIN (FRAGMENT).  
 OS *Mus musculus* (Mouse).  
 OC Eukaryota; Metazoa;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8652293; PubMed=3755150;  
 RA Ayane M., Mengle-Gaw L., McDewitt H.O., Benoist C., Mathis D.;  
 RT "E alpha u and E beta u chain association: where lies the anomaly?";  
 RL J. Immunol. 137:948-951(1986).  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M12818; AAA39638.1; .  
 DR HSP; P01903; ISEB.  
 DR InterPro; IPR000495; .  
 DR InterPro; IPR001003; .  
 DR InterPro; IPR003006; .  
 DR Pfam; PF00993; MHC\_II\_alpha; 1.  
 DR Pfam; PF00047; Ig; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW MHC II; Transmembrane; Glycoprotein.  
 FT NON\_TER 1  
 FT DOMAIN <1 83 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 84 177 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 178 190 CONNECTING PEPTIDE.  
 FT TRANSMEM 191 216

FT DOMAIN 217 229 CYTOPLASMIC TAIL.  
 FT DISULFID 106 162 BY SIMILARITY.  
 FT CARBOHYD 117 117 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 229 AA; 26210 MW; 1E5A6463F1FE64D3 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 54;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQGL 1115  
 DB 53 FEAQGL 59  
 |||||

RESULT 27  
 RS4E\_ARCFU  
 ID RS4E\_ARCFU STANDARD; PRT; 235 AA.  
 AC O28366;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 30S RIBOSOMAL PROTEIN S4E.  
 GN RPS4E OR AFL1913.  
 OS *Archaeoglobus fulgidus*.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 OC *Archaeoglobus*.  
 OX NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kervatage A.R., Graham D.E., Kyriades N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterback T.,  
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon *Archaeoglobus fulgidus*.";  
 RL Nature 390:364-370(1997).  
 CC -!- SIMILARITY: BELONGS TO THE S4E FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE000971; AAB89340.1; .  
 DR TIGR; AF1913; .  
 DR InterPro; IPR000876; .  
 DR Pfam; PF00900; Ribosomal\_S4e; 1.  
 DR PROSITE; PS00528; RIBOSOMAL\_S4E; 1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 235 AA; 26439 MW; 034C049FFEF9E1DF CRC64;

Query Match 0.5%; Score 7; DB 1; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 686 LVRRGV 692  
 DB 125 LVRRGV 131  
 |||||

RESULT 28  
BCL2\_MOUSE  
ID BCL2\_MOUSE STANDARD; PRT; 236 AA.  
AC P10417; P10418;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE APOPTOSIS REGULATOR BCL-2.  
GN BCL2 OR BCL-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=Liver;  
RX MEDLINE=87187643; PubMed=3032455;  
RA Negri M., Silini E., Kozak C., Tsujimoto Y., Croce C.M.;  
RT "Molecular analysis of mbcl-2: structure and expression of the murine  
RT gene homologous to the human gene involved in follicular lymphoma.";  
RL Cell 49:455-463(1987).  
RN [2]  
RP REVISIONS TO 221-222.  
RX MEDLINE=92375724; PubMed=1508712;  
RA Eguchi Y., Ewert D.L., Tsujimoto Y.;  
RT "Isolation and characterization of the chicken bcl-2 gene: expression  
RT in a variety of tissues including lymphoid and neuronal organs in  
RT adult and embryo.";  
RL Nucleic Acids Res. 20:4187-4192(1992).  
CC -1- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE  
CC ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF  
CC VARIOUS STIMULI INDUCING CELLULAR DEATH. BCL2 BLOCKS APOPTOSIS  
CC BECAUSE IT INTERFERES WITH THE ACTIVATION OF CASPASES BY  
CC PREVENTING THE RELEASE OF CYTOCHROME C. MIGHT FUNCTION IN AN  
CC ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES OF FREE RADICAL  
CC GENERATION SUCH AS MITOCHONDRIA.  
CC -1- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BAX AND  
CC BAK PROTEINS, AND WITH BCL-X(S). HETERODIMERIZATION WITH BAX  
CC REQUIRES INTACT BH1 AND BH2 DOMAINS, AND IS NECESSARY FOR ANTI-  
CC APOPTOTIC ACTIVITY (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER AND OUTER MEMBRANES, AS  
CC WELL AS NUCLEAR ENVELOPE AND ENDOPLASMIC RETICULUM.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA; ARE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.  
CC -1- DOMAIN: BH4 DOMAIN SEEMS TO BE INVOLVED IN THE ANTI-APOPTOTIC  
CC FUNCTION.  
CC -1- SIMILARITY: CONTAINS A BCL-2 HOMOLGY DOMAIN 1 (BH1).  
CC -1- SIMILARITY: CONTAINS A BCL-2 HOMOLGY DOMAIN 2 (BH2).  
CC -1- SIMILARITY: CONTAINS A BCL-2 HOMOLGY DOMAIN 3 (BH3).  
CC -1- SIMILARITY: CONTAINS A BCL-2 HOMOLGY DOMAIN 4 (BH4).  
CC -1- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
CC -----  
DR EMBL; L31532; AAA37282.1; -  
DR EMBL; M16506; AAA37282.1; JOINED.  
DR EMBL; M16506; AAA37281.1; -  
DR PIR; A25960; TVMSA1.  
DR PIR; B25960; TVMSB1.  
DR PIR; E37332; E37332.  
DR HSP; Q07817; IMAZ.  
DR MGD; MGI:88138; Bcl2.  
DR InterPro; IPR000712; -  
DR InterPro; IPR002475; -  
DR InterPro; IPR003093; -

DR Pfam; PF00452; Bcl-2; 1.  
DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
DR PROSITE; PS01080; BH1; 1.  
DR PROSITE; PS01258; BH2; 1.  
DR PROSITE; PS01259; BH3; 1.  
DR PROSITE; PS01260; BH4\_1; 1.  
DR PROSITE; PS00063; BH4\_2; 1.  
KW Apoptosis; Alternative splicing; Transmembrane; Mitochondrion.  
FT DOMAIN 10 30  
FT BH4.  
FT DOMAIN 90 104  
FT BH3.  
FT DOMAIN 133 152  
FT BH1.  
FT DOMAIN 184 199  
FT BH2.  
FT TRANSMEM 209 230  
FT POTENTIAL.  
FT VARSPLIC 193 236  
FT DAFVLYGFSMRPLDFSWLSKTLISLALVGACITLGAYL  
FT GHK -> VGACIVE (IN ISOFORM BETA).  
SQ SEQUENCE 236 AA; 26425 MW; AA85EF6B0766BE0A CRC64;  
  
Query Match 0.5%; Score 7; DB 1; Length 236;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 16 SLALVGA 22  
DB 219 SLALVGA 225  
|||||||  
  
RESULT 29  
BCL2\_MOUSE  
ID BCL2\_MOUSE STANDARD; PRT; 236 AA.  
AC P4950; O62837; O64032;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE APOPTOSIS REGULATOR BCL-2.  
GN BCL2 OR BCL-2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=94193015; PubMed=8144041;  
RA Sato T., Irie S., Krajewski S., Reed J.C.;  
RT "Cloning and sequencing of a cDNA encoding the rat Bcl-2 protein.";  
RL Gene 140:291-292(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Ovary;  
RX MEDLINE=95129487; PubMed=7828536;  
RA Tilly J.L., Tilly K.I., Kenton M.L., Johnson A.L.;  
RT "Expression of members of the bcl-2 gene family in the immature rat  
RT ovary: equine chorionic gonadotropin-mediated inhibition of granulosa  
RT cell apoptosis is associated with decreased bax and constitutive  
RT bcl-2 and bcl-x long messenger ribonucleic acid levels.";  
RL Endocrinology 136:232-241(1995).  
RN [3]  
RP SEQUENCE OF 19-172 FROM N.A.  
RX MEDLINE=95059917; PubMed=7969891;  
RA Castren E., Ohga Y., Bezaghi M.P., Tzimagiorgis G., Thoenen H.,  
RA Lindholm D.;  
RT "bcl-2 messenger RNA is localized in neurons of the developing and  
RT adult rat brain.";  
RL Neuroscience 61:165-177(1994).  
CC -1- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE  
CC ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF  
CC VARIOUS STIMULI INDUCING CELLULAR DEATH. BCL2 BLOCKS APOPTOSIS  
CC BECAUSE IT INTERFERES WITH THE ACTIVATION OF CASPASES BY  
CC PREVENTING THE RELEASE OF CYTOCHROME C. MIGHT FUNCTION IN AN  
CC ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES OF FREE RADICAL  
CC GENERATION SUCH AS MITOCHONDRIA.  
CC -1- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BAX AND



CC BAK PROTEINS, AND WITH BCL-X(S). HETERODIMERIZATION WITH BAX  
 CC REQUIRES INTACT BH1 AND BH2 DOMAINS, AND IS NECESSARY FOR ANTI-  
 CC APOPTOTIC ACTIVITY (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER AND OUTER MEMBRANES, AS  
 CC WELL AS NUCLEAR ENVELOPE AND ENDOPLASMIC RETICULUM.  
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES, WITH  
 CC HIGHEST LEVELS IN REPRODUCTIVE TISSUES. IN THE ADULT BRAIN,  
 CC EXPRESSION IS LOCALIZED IN MITRAL CELLS OF THE OLFACTORY BULB,  
 CC GRANULE AND PYRAMIDAL NEURONS OF HIPPOCAMPUS, PONTINE NUCLEI,  
 CC CEREBELLAR GRANULE NEURONS, AND IN EPENDYMAL CELLS. IN PRENATAL  
 CC BRAIN, EXPRESSION IS HIGHER AND LOCALIZED IN THE NEUROEPITHELIUM  
 CC AND IN THE CORTICAL PLATE.  
 CC -!- DOMAIN: BH4 DOMAIN SEEMS TO BE INVOLVED IN THE ANTI-APOPTOTIC  
 CC FUNCTION.  
 CC -!- SIMILARITY: CONTAINS A BCL-2 HOMOLOGY DOMAIN 1 (BH1).  
 CC -!- SIMILARITY: CONTAINS A BCL-2 HOMOLOGY DOMAIN 2 (BH2).  
 CC -!- SIMILARITY: CONTAINS A BCL-2 HOMOLOGY DOMAIN 3 (BH3).  
 CC -!- SIMILARITY: CONTAINS A BCL-2 HOMOLOGY DOMAIN 4 (BH4).  
 CC -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; L14680; AAA53662.1; -  
 CC EMBL; U34964; AAA77687.1; -  
 CC EMBL; S74122; -; NOT\_ANNOTATED\_CDS.  
 CC HSSP; P53563; IAF3.  
 CC InterPro; IPR000712; -  
 CC InterPro; IPR002475; -  
 CC InterPro; IPR003093; -  
 CC Pfam; PF00452; BCL-2; 1.  
 CC PROSITE; PS00062; BCL2\_FAMILY; 1.  
 CC PROSITE; PS01080; BH1; 1.  
 CC PROSITE; PS01258; BH2; 1.  
 CC PROSITE; PS01259; BH3; 1.  
 CC PROSITE; PS01260; BH4\_1; 1.  
 CC PROSITE; PS00063; BH4\_2; 1.  
 CC Apoptosis; Alternative splicing; Transmembrane; Mitochondrion.  
 CC  
 CC FT DOMAIN 10 30 BH4.  
 CC FT DOMAIN 90 104 BH3.  
 CC FT DOMAIN 133 152 BH1.  
 CC FT DOMAIN 184 199 BH2.  
 CC FT TRANSMEM 209 230 POTENTIAL.  
 CC FT CONFLICT 42 42 A -> R (IN REF. 2).  
 CC FT CONFLICT 157 157 E -> G (IN REF. 1).  
 CC FT CONFLICT 164 164 S -> Y (IN REF. 2).  
 CC FT CONFLICT 212 212 L -> Q (IN REF. 2).  
 CC SEQUENCE 236 AA; 26622 MW; E7688CB9071A872A CRC64;  
 CC  
 CC Query Match 0.5%; Score 7; DB 1; Length 236;  
 CC Best Local Similarity 100.08; Pred. No. 55;  
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC Qy 16 SLALVGA 22  
 CC |||||  
 CC Db 219 SLALVGA 225  
 CC  
 CC RESULT 30  
 CC BCL2\_HUMAN  
 CC ID BCL2\_HUMAN STANDARD; PRT; 239 AA.  
 CC AC P10415; P10416; Q16197; Q13842;  
 CC DT 01-MAR-1989 (rel. 10; Created)  
 CC DT 01-APR-1993 (rel. 25; Last sequence update)  
 CC DT 01-OCT-2000 (rel. 40; Last annotation update)

DE APOPTOSIS REGULATOR BCL-2.  
 GN BCL2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86259760; PubMed=3523487;  
 RA Tsujimoto Y., Croce C.M.;  
 RT "Analysis of the structure, transcripts, and protein products of  
 RT bcl-2, the gene involved in human follicular lymphoma.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5214-5218(1986).  
 RN [2]  
 RP REVISIONS TO 96; 110 AND 237.  
 RX MEDLINE=92375724; PubMed=1508712;  
 RA Eguchi Y., Ewert D.L., Tsujimoto Y.;  
 RT "Isolation and characterization of the chicken bcl-2 gene: expression  
 RT in a variety of tissues including lymphoid and neuronal organs in  
 RT adult and embryo.";  
 RL Nucleic Acids Res. 20:4187-4192(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87002488; PubMed=2875799;  
 RA Cleary M.L., Smith S.D., Sklar J.;  
 RT "Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-  
 RT 2/immunoglobulin transcript resulting from the t(14;18)  
 RT translocation.";  
 RL Cell 47:19-28(1986).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88196071; PubMed=2834197;  
 RA Seto M., Jaeger U., Hockett R.D., Graninger W., Bennett S.,  
 RA Goldman P., Korsmeyer S.J.;  
 RT "Alternative promoters and exons, somatic mutation and deregulation  
 RT of the Bcl-2-Ig fusion gene in lymphoma.";  
 RL EMBO J. 7:123-131(1988).  
 RN [5]  
 RP SEQUENCE OF 1-131 FROM N.A., AND VARIANTS NON-HODGKINS-LYMPHOMA.  
 RX MEDLINE=92096610; PubMed=1339299;  
 RA Tanaka S., Louie D.C., Kant J.A., Reed J.C.;  
 RT "Frequent incidence of somatic mutations in translocated BCL2  
 RT oncogenes of non-Hodgkin's lymphomas.";  
 RL Blood 79:229-237(1992).  
 RN [6]  
 RP SUBCELLULAR LOCATION.  
 RX MEDLINE=91066924; PubMed=2250705;  
 RA Hockenbery D., Nunez G., Millman C., Schreiber R.D., Korsmeyer S.J.;  
 RT "Bcl-2 is an inner mitochondrial membrane protein that blocks  
 RT programmed cell death.";  
 RL Nature 348:334-336(1990).  
 RN [7]  
 RP MUTAGENESIS.  
 RX MEDLINE=94239528; PubMed=8183370;  
 RA Yin X.-M., Oltvai Z.N., Korsmeyer J.;  
 RT "BH1 and BH2 domains of Bcl-2 are required for inhibition of  
 RT apoptosis and heterodimerization with Bax.";  
 RL Nature 369:321-323(1994).  
 CC -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE  
 CC ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF  
 CC VARIOUS STIMULI INDUCING CELLULAR DEATH. BCL2 BLOCKS APOPTOSIS  
 CC BECAUSE IT INTERFERES WITH THE ACTIVATION OF CASPASES BY  
 CC PREVENTING THE RELEASE OF CYTOCHROME C. MIGHT FUNCTION IN AN  
 CC ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES OF FREE RADICAL  
 CC GENERATION SUCH AS MITOCHONDRIA.  
 CC -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BAX AND  
 CC BAK PROTEINS, AND WITH BCL-X(S). HETERODIMERIZATION WITH BAX  
 CC REQUIRES INTACT BH1 AND BH2 DOMAINS, AND IS NECESSARY FOR ANTI-  
 CC APOPTOTIC ACTIVITY.  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER AND OUTER MEMBRANES, AS  
 CC WELL AS NUCLEAR ENVELOPE AND ENDOPLASMIC RETICULUM.  
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA;  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.





HA2R\_HUMAN  
ID HA2R\_HUMAN STANDARD; PRT; 254 AA.  
AC P01903;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR ALPHA CHAIN PRECURSOR.  
GN HLA-DR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83221632; PubMed=6304715;  
RA Das H.K., Lawrence S.K., Weissman S.M.;  
RT "Structure and nucleotide sequence of the heavy chain gene of  
RT HLA-DR.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:3543-3547(1983).  
RN [2]  
RP REVISIONS.  
RA Das H.K., Lawrence S.K., Weissman S.M.;  
RL Proc. Natl. Acad. Sci. U.S.A. 80:7024-7024(1983).  
RN [3]  
RP SEQUENCE OF 26-204.  
RX MEDLINE=82263347; PubMed=6955253;  
RA Yang C.-Y., Kratzin H., Gotz H., Thinnies F.P., Kruse T., Egert G.,  
RA Pauly E., Kolbel S., Wernet P., Hilschmann N.;  
RT "Primary structure of class II human histocompatibility antigens. 2nd  
RT Communication. Amino acid sequence of the N-terminal 179 residues of  
RT the alpha-chain of an HLA-Dw2/DR2 alloantigen.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 363:671-676(1982).  
RN [4]  
RP SEQUENCE OF 26-60, AND SEQUENCE OF 32-202 AND 204-254 FROM N.A.  
RX MEDLINE=83025073; PubMed=6812963;  
RA Larhammar D., Gustafsson K., Claesson L., Bill P., Wiman K.,  
RA Schenning L., Sundelin J., Widmark E., Peterson P.A., Rask L.;  
RT "Alpha chain of HLA-DR transplantation antigens is a member of the  
RT same protein superfamily as the immunoglobulins.";  
RL Cell 30:153-161(1982).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84057142; PubMed=6416803;  
RA Kajimura Y., Toyoda H., Sato M., Miyakoshi S., Kaplan S.A., Ike Y.,  
RA Goyert S.M., Silver J., Hawke D., Shively J.E., Suggs S.V.,  
RA Wallace R.B., Itakura K.;  
RT "Cloning the heavy chain of human HLA-DR antigen using synthetic  
RT oligodeoxyribonucleotides as hybridization probes.";  
RL DNA 2:1175-1182(1983).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84169507; PubMed=6324094;  
RA Schamboeck A., Korman A.J., Kamb A., Strominger J.L.;  
RT "Organization of the transcriptional unit of a human class II  
RT histocompatibility antigen: HLA-DR heavy chain.";  
RL Nucleic Acids Res. 11:8663-8675(1983).  
RN [7]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83013020; PubMed=6811954;  
RA Lee J.S., Trowsdale J., Travers P.J., Carey J., Grosveld F.,  
RA Jenkins J., Bodmer W.F.;  
RT "Sequence of an HLA-DR alpha-chain cDNA clone and intron-exon  
RT organization of the corresponding gene.";  
RL Nature 299:750-752(1982).  
RN [8]  
RP SEQUENCE OF 29-254 FROM N.A.  
RX MEDLINE=83299916; PubMed=6821129;  
RA Korman A.J., Auffray C., Schamboeck A., Strominger J.L.;  
RT "The amino acid sequence and gene organization of the heavy chain of  
RT the HLA-DR antigen: homology to immunoglobulins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:6013-6017(1982).  
RN [9]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91010755; PubMed=2212658;  
RA Koppelman B., Cresswell P.;  
RT "Rapid nonlysosomal degradation of assembled HLA class II  
RT glycoproteins incorporating a mutant DR alpha-chain.";  
RL J. Immunol. 145:2730-2736(1990).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 28-207.  
RX MEDLINE=94195388; PubMed=8145819;  
RA Stern L.J., Brown J.H., Jardetzky T.J., Gorga J.C., Urban R.G.,  
RA Strominger J.L., Wiley D.C.;  
RT "Crystal structure of the human class II MHC protein HLA-DR1  
RT complexed with an influenza virus peptide.";  
RL Nature 368:215-221(1994).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 28-207.  
RX MEDLINE=93302847; PubMed=8316295;  
RA Brown J.H., Jardetzky T.S., Gorga J.C., Stern L.J., Urban R.G.,  
RA Strominger J.L., Wiley D.C.;  
RT "Three-dimensional structure of the human class II histocompatibility  
RT antigen HLA-DR1.";  
RL Nature 364:33-39(1993).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF COMPLEX WITH SEB.  
RX MEDLINE=94203282; PubMed=8152483;  
RA Jardetzky T.S., Brown J.H., Gorga J.C., Stern L.J., Urban R.G.,  
RA Chi Y.L., Stauffer C., Strominger J.L., Wiley D.C.;  
RT "Three-dimensional structure of a human class II histocompatibility  
RT molecule complexed with superantigen.";  
RL Nature 368:711-718(1994).  
RN [13]  
RP X-RAY CRYSTALLOGRAPHY (2.75 ANGSTROMS).  
RX MEDLINE=96085023; PubMed=7477400;  
RA Ghosh P., Amaya M., Mellins E., Wiley D.C.;  
RT "The structure of an intermediate in class II MHC maturation: CLIP  
RT bound to HLA-DR3.";  
RL Nature 378:457-462(1995).  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF COMPLEX WITH COLLAGEN.  
RX MEDLINE=98014591; PubMed=9354468;  
RA Dessen A., Lawrence C.M., Cupo S., Zaller D.M., Wiley D.C.;  
RT "X-ray crystal structure of HLA-DR4 (DRA\*0101, DRB1\*0401) complexed  
RT with a peptide from human collagen II.";  
RL Immunity 7:473-481(1997).  
RN [15]  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -  
RP the European Bioinformatics Institute. There are no restrictions on its  
RP use by non-profit institutions as long as its content is in no way  
RP modified and this statement is not removed. Usage by and for commercial  
RP entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
RP or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RN [16]  
RP EMBL; V00523; CAA23782.1;  
DR EMBL; J00204; AAA36302.1;  
DR EMBL; J00203; AAA36302.1; JOINED;  
DR EMBL; X00274; CAA25076.1; ALT\_INIT.  
DR EMBL; J00194; AAA36275.1;  
DR EMBL; K01171; AAA59785.1;  
DR EMBL; M60334; AAA59783.1;  
DR PIR; A02206; HLHDA.  
DR PDB; 1DLH; 22-JUN-94.  
DR PDB; 1SEB; 20-JUN-96.  
DR PDB; 2SEB; 28-JAN-98.  
DR PDB; 1AQD; 28-JAN-98.  
DR PDB; 1A6A; 27-MAY-98.  
DR MIM; 142860;  
DR InterPro; IPR000495;  
DR InterPro; IPR001003;  
DR InterPro; IPR003006;  
DR Pfam; PF00993; MHC\_II\_alpha; 1.  
DR Pfam; PF00047; Ig; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
MHC II; Transmembrane; Glycoprotein; Signal; 3D-structure.

```

FT SIGNAL 1 25 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN,
FT CHAIN 26 254 DR ALPHA CHAIN.
FT DOMAIN 26 109 EXTRACELLULAR ALPHA-1.
FT DOMAIN 110 203 EXTRACELLULAR ALPHA-2.
FT DOMAIN 204 213 CONNECTING PEPTIDE.
FT TRANSMEM 217 239 CYTOPLASMIC.
FT DOMAIN 240 254 BY SIMILARITY.
FT DISULFID 132 188 N-LINKED (GLCNAC. . .).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .).
FT VARIANT 143 143 V -> L.
FT CONFLICT 28 29 /FTIG=VAR_004399.
FT CONFLICT 33 33 EE -> AD (IN REF. 4).
FT CONFLICT 34 35 I -> T (IN REF. 4).
FT CONFLICT 48 48 QA -> YP (IN REF. 4).
FT CONFLICT 54 54 M -> Q (IN REF. 4).
FT CONFLICT 67 67 D -> T (IN REF. 4).
FT CONFLICT 149 149 V -> A (IN REF. 3).
FT CONFLICT 149 149 N -> E (IN REF. 3).
SQ SEQUENCE 254 AA; 28607 MW; 3CDICDBA952B2350 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 254;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAGAL 1115
Db 79 FEAQAGAL 85

RESULT 33
HA21_MOUSE STANDARD; PRT; 255 AA.
AC P01904;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE H-2 CLASS II HISTOCOMPATIBILITY ANTIGEN, E-D ALPHA CHAIN PRECURSOR
DE (H2-IE-ALPHA).
GN H2-EA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RX SEQUENCE FROM N.A.
RA Hyldig-Nielsen J.J., Schenning L., Hammerling U., Widmark E.,
RA Heldin E., Lind P., Servenius B., Lund T., Flavell R., Lee J.S.,
RA Trowsdale J., Schreier P.H., Zablitzy F., Larhammar D.,
RA Peterson P.A., Rask L.;
RA "The complete nucleotide sequence of the I-E alpha d immune response
gene.";
RL Nucleic Acids Res. 11:5055-5071(1983).
RN [2]
RX SEQUENCE FROM N.A.
RA Larhammar D., Andersson G., Andersson M., Bill P., Boehme J.,
RA Claesson L., Denaro M., Emmoth E., Gustafsson K., Hammarling U.,
RA Heldin E., Hyldig-Nielsen J.J., Lind P., Schenning L., Servenius B.,
RA Widmark E., Rask L., Peterson P.A.;
RT "Molecular analysis of human class II transplantation antigens and
their genes.";
RL Hum. Immunol. 8:95-103(1983).
RN [3]
RX SEQUENCE OF 29-255 FROM N.A.
RX MEDLINE=83067428; PubMed=6815800;
RA McNicholas J., Steinmetz M., Hunkapiller T., Jones P., Hood L.E.;
RT "DNA sequence of the gene encoding the E alpha Ia polypeptide of the
BALB/c mouse.";
RL Science 218:1229-1232(1982).

```

```

RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=83169693; PubMed=6300851;
RA Benoist C.O., Mathis D.J., Kanter M.R., Williams V.E., McDewitt H.O.;
RT "The murine Ia alpha chains, E alpha and A alpha, show a surprising
degree of sequence homology.";
RN [5]
RL Proc. Natl. Acad. Sci. U.S.A. 80:534-538(1983).
RP SEQUENCE FROM N.A.
RX MEDLINE=83155651; PubMed=6403249;
RA Mathis D.J., Benoist C.O., Williams V.E. II, Kanter M.R.,
RA McDewitt H.O.;
RT "The murine E alpha immune response gene.";
RL Cell 32:745-754(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K00971; AAA98624.1; -
DR PIR; A02207; HLMSED.
DR PIR; A21217; A21217.
DR HSPP; P01903; 25EB.
DR MGD; MGI:95900; H2-Ea.
DR InterPro; IPR000495; -
DR InterPro; IPR001003; -
DR InterPro; IPR003006; -
DR Pfam; PF00993; MHC_II_alpha; 1.
DR Pfam; PF00047; Ig; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC II; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 255 H-2 CLASS II HISTOCOMPATIBILITY ANTIGEN,
FT DOMAIN 26 109 E-D ALPHA CHAIN.
FT DOMAIN 110 203 EXTRACELLULAR ALPHA-1.
FT DOMAIN 204 216 EXTRACELLULAR ALPHA-2.
FT TRANSMEM 217 242 CONNECTING PEPTIDE.
FT DOMAIN 243 255 CYTOPLASMIC TAIL.
FT DISULFID 132 188 BY SIMILARITY.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 155 155 E -> T (IN REF. 3).
FT CONFLICT 202 202 T -> D (IN REF. 3).
FT CONFLICT 202 202 T -> H (IN REF. 4 AND 5).
FT CONFLICT 219 219 M -> V (IN REF. 4 AND 5).
FT CONFLICT 239 239 M -> A (IN REF. 2).
SQ SEQUENCE 255 AA; 29116 MW; 0852EA3AA4EE2674 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAGAL 1115
Db 79 FEAQAGAL 85

RESULT 34
HA22_MOUSE STANDARD; PRT; 255 AA.
ID HA22_MOUSE
AC P04224;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE H-2 CLASS II HISTOCOMPATIBILITY ANTIGEN, E-K ALPHA CHAIN PRECURSOR.
DE Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83169693; PubMed=6300851;
RA Benoist C.O., Mathis D.J., Kanter M.R., Williams V.E., McDevitt H.O.;
RT "The murine Ia alpha chains, E alpha and A alpha, show a surprising
RL degree of sequence homology.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:534-538(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83155651; PubMed=6403249;
RA Mathis D.J., Benoist C.O., Williams V.E., Kanter M.R., McDevitt H.O.;
RT "The murine E alpha immune response gene.";
RL Cell 32:745-754(1983).
RN [3]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00833; CAA24216.2;
DR EMBL; V00834; CAA24216.2; JOINED.
DR PIR; A02208; HLMSEA.
DR PIR; A21938; A21938.
DR HSSP; P01903; ISEB.
DR InterPro; IPR000495;
DR InterPro; IPR001003;
DR InterPro; IPR003006;
DR Pfam; PF00993; MHC_Ii_alpha; 1.
DR Pfam; PF00047; Ig; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC II; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 255
FT -----
FT H-2 CLASS II HISTOCOMPATIBILITY ANTIGEN,
FT E-K ALPHA CHAIN.
FT DOMAIN 26 109
FT DOMAIN 110 203
FT DOMAIN 204 216
FT TRANSMEM 217 242
FT DOMAIN 243 255
FT DISULFID 132 188
FT CARBOHYD 143 143
FT SEQUENCE 255 AA; 29120 MW; 144EEA3AB1EF3724 CRC64;
SQ
Query Match 0.5%; Score 7; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1109 FEAQGL 1115
Db 79 FEAQGL 85
|||||
RESULT 35
HMX2_CHICK
ID HMX2_CHICK STANDARD; PRT; 259 AA.
AC P28362;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HMOBOX PROTEIN MSX-2 (CHOX-8) (GHOX-8).
GN MSX2 OR CHOX-8.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RX MEDLINE=92001542; PubMed=1680378;
RA Coelho C.N., Sumoy L., Rodgers B.J., Davidson D.R., Hill R.E.,
RA Upholt W.B., Koshier R.A.;
RT "Expression of the chicken homeobox-containing gene GHox-8 during
RL embryonic chick limb development.";
RL Mech. Dev. 34:143-154(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92146256; PubMed=1685987;
RA Yokouchi Y., Ohsugi K., Sasaki H., Kuroiwa A.;
RT "Chicken homeobox gene Msx-1: structure, expression in limb buds and
RL effect of retinoic acid.";
RL Development 113:431-444(1991).
RN [3]
RP SEQUENCE OF 134-259 FROM N.A.
RX MEDLINE=92090717; PubMed=1684333;
RA Robert B., Lyons B., Simandl B.K., Kuroiwa A., Buckingham M.;
RT "The apical ectodermal ridge regulates Hox-7 and Hox-8 gene
RL expression in developing chick limb buds.";
RT Genes Dev. 5:2363-2374(1991).
CC -!- FUNCTION: MORPHOGENETIC ROLE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE MSH FAMILY OF HMOBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S64478; AAB20018.1;
DR EMBL; X62097; CAA44007.1;
DR EMBL; X62541; CAA44425.1;
DR PIR; B41635; B41635.
DR HSSP; P22808; IVND.
DR TRANSFAC; T02074;
DR InterPro; IPR001356;
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HMOBOX.
DR PROSITE; PS00027; HMOBOX_1; 1.
DR PROSITE; PS00071; HMOBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DNA_BIND 134 193 HMOBOX.
FT CONFLICT 58 58 S -> N (IN REF. 2).
FT SEQUENCE 259 AA; 28235 MW; 375405ED73E59786 CRC64;
SQ
Query Match 0.5%; Score 7; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1046 ASLYGTS 1052
Db 225 ASLYGTS 231
|||||
RESULT 36
HMX2_COTJA
ID HMX2_COTJA STANDARD; PRT; 259 AA.
AC P23410;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HMOBOX PROTEIN MSX-2 (MSX-1) (QUOX-7).
GN MSX2 OR MSX1 OR QUOX-7.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
```

RP SEQUENCE FROM N.A.  
RX MEDLINE-91017530; PubMed-1977161;  
RA Takahashi Y., le Douarin N.M.;  
RT "CDNA cloning of a quail homeobox gene and its expression in neural  
crest-derived mesenchyme and lateral plate mesoderm.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:7482-7486(1990).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN NEURAL CREST-DERIVED MESENCHYME  
CC AND LATERAL PLATE MESODERM.  
CC -1- SIMILARITY: BELONGS TO THE MSH FAMILY OF HOMEBOX PROTEINS.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M57611; AAA63459.1; -;  
DR PIR; A38284; A38284.  
DR HSP; P22808; IYND.  
DR TRANSFAC; T02077; -;  
DR InterPro; IPR001356; -;  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
DR Homeobox; DNA-binding; Developmental protein; Nuclear protein.  
FT DNA\_BIND 134 193 HOMEBOX.  
SQ SEQUENCE 259 AA; 28243 MW; 0CDF2512FDCC9A79 CRC64;  
  
Query Match 0.5%; Score 7; DB 1; Length 259;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1046 ASLYGTS 1052  
Db 225 ASLYGTS 231  
|||||  
  
RESULT 37  
CTRB\_HUMAN STANDARD; PRT; 263 AA.  
AC P17538;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE CHYMOTRYPSINOGEN B PRECURSOR (EC 3.4.21.1).  
GN CTRB1 OR CTRB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RX MEDLINE-89134264; PubMed-2917002;  
RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,  
RA Mori T., Matsubara K.;  
RT "Molecular cloning and nucleotide sequence of human pancreatic  
prechymotrypsinogen cDNA.";  
RL Biochem. Biophys. Res. Commun. 158:569-575(1989).  
CC -1- SUBCELLULAR LOCATION: PREFERENTIAL CLEAVAGE: TYR-, TRP-, PHE-, LEU-.  
CC -1- TISSUE SPECIFICITY: EXTRACELLULAR.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M24400; AAA52128.1; -;  
DR PIR; A31299; A31299.  
DR HSP; P00766; 1CHG.  
DR MEROPS; S01.152; -;  
DR MIM; 118890; -;  
DR InterPro; IPR001254; -;  
DR InterPro; IPR001314; -;  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 263 CHYMOTRYPSINOGEN B.  
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.  
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.  
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.  
FT ACT\_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 19 140 BY SIMILARITY.  
FT DISULFID 60 76 BY SIMILARITY.  
FT DISULFID 154 219 BY SIMILARITY.  
FT DISULFID 186 200 BY SIMILARITY.  
FT DISULFID 209 238 BY SIMILARITY.  
SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A490B8701 CRC64;  
  
Query Match 0.5%; Score 7; DB 1; Length 263;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 71 ANKTPDK 77  
Db 166 ANKTPDK 172  
|||||  
  
RESULT 38  
TVSY\_MYCTU STANDARD; PRT; 263 AA.  
AC O33306;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS) (TSASE).  
GN THYA OR RV2764C OR MTV002.29C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE-98295987; PubMed-9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,  
RA Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).  
CC -1- FUNCTION: PROVIDES THE SOLE DE NOVO SOURCE OF DTMP FOR DANA  
CC BIOSYNTHESIS (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + DUMP =  
CC DIHYDROFOLATE + DTMP.

CC -1- PATHWAY: DEOXYRIBONUCLEOTIDE BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE THYMIDYLATE SYNTHASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL008967; CAAL5560.1; -  
 CC HSSP; P00470; 1AN5.  
 CC TubercuList; Rv2764c; -  
 CC InterPro; IPR000398; -  
 CC Pfam; PF00303; thymidylat\_synt; 1.  
 CC PRINTS; PS00108; THYMIDSYNTHASE.  
 CC PROSITE; PS00091; THYMIDYLATE\_SYNTHASE; 1.  
 CC Transferase; Methyltransferase; Nucleotide biosynthesis.  
 CC ACT\_SITE 146 146 BY SIMILARITY.  
 CC SEQUENCE 263 AA; 29853 MW; DF48279E3E9EA5F3 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 263;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 590 TGTSLF 596  
 |||||  
 DB 24 TGTSLF 30

RESULT 39  
 LECA\_SARPE  
 ID LECA\_SARPE STANDARD; PRT; 283 AA.  
 AC P05047;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE LECTIN, ALPHA SUBUNIT PRECURSOR.  
 OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestroidea; Sarcophagidae; Sarcophaga.  
 ON NCBI\_TaxID=7386;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=86008294; PubMed=2413021;  
 RA Takahashi H., Komano H., Kawaguchi N., Kitamura N., Nakanishi S.,  
 RA Natori S.;  
 RT "Cloning and sequencing of cDNA of Sarcophaga peregrina humoral  
 RT lectin induced on injury of the body wall.";  
 RL J. Biol. Chem. 260:12228-12233(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94139722; PubMed=8307011;  
 RA Matsui M., Kobayashi A., Kubo T., Natori S.;  
 RA "Purification and characterization of ATP, a novel protein that  
 RT binds to A/T stretches in three segments of the Sarcophaga lectin  
 RT gene.";  
 RL Eur. J. Biochem. 219:449-454(1994).  
 RN [3]  
 RP SEQUENCE OF 1-38 FROM N.A.  
 RX MEDLINE=90089397; PubMed=2480809;  
 RA Kobayashi A., Hirai H., Kubo T., Veno K., Nakanishi Y., Natori S.;  
 RA "Cloning and in vitro transcription of the Sarcophaga lectin gene.";  
 RL Biochim. Biophys. Acta 1009:244-250(1989).  
 CC -1- FUNCTION: ROLE IN THE DEFENCE SYSTEM OF THE ORGANISM AGAINST  
 CC MICROORGANISMS. THIS LECTIN BINDS GALACTOSE.  
 CC -1- INDUCTION: THIS HUMORAL LECTIN IS INDUCED ON INJURY OF THE BODY  
 CC WALL.  
 CC -1- SIMILARITY: TO OTHER MEMBERS OF THE C-TYPE LECTIN FAMILY.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M11673; AAA29983.1; -  
 CC EMBL; D14870; BAA03586.1; -  
 CC EMBL; X16659; CAA34645.1; -  
 CC PIR; A25736; LMFHLS.  
 CC PIR; S07759; S07759.  
 CC PIR; S41119; S41119.  
 CC InterPro; IPR001304; -  
 CC Pfam; PF00059; lectin\_c; 1.  
 CC PROSITE; PS00615; C-TYPE\_LECTIN\_1; 1.  
 CC PROSITE; PS00041; C-TYPE\_LECTIN\_2; 1.  
 KW Lectin; Signal.  
 FT SIGNAL 1 23  
 FT CHAIN 24 283 LECTIN, ALPHA SUBUNIT.  
 FT DOMAIN 51 159 C-TYPE LECTIN (SHORT FORM).  
 FT DISULFID 53 157 BY SIMILARITY.  
 FT DISULFID 132 149 BY SIMILARITY.  
 SQ SEQUENCE 283 AA; 32991 MW; F324BF1A1140B3AC CRC64;

Query Match 0.5%; Score 7; DB 1; Length 283;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 NNPDNYK 807  
 |||||  
 DB 121 NNPDNYK 127

RESULT 40

ALF\_STRPN  
 ID ALF\_STRPN STANDARD; PRT; 293 AA.  
 AC O65944;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE FRUCTOSE-BISPHOSPHATE ALDOLASE (EC 4.1.2.13).  
 GN FBA.  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 ON NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R6;  
 RA Jado I., Casal J., Fenoll A., Perez A.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE = GLYCERONE-  
 CC PHOSPHATE + D-GLYCERALDEHYDE 3-PHOSPHATE.  
 CC -1- COFACTOR: ZINC (BY SIMILARITY).  
 CC -1- PATHWAY: SIXTH STEP IN GLYCOLYSIS.  
 CC -1- SIMILARITY: BELONGS TO CLASS II FRUCTOSE-BISPHOSPHATE ALDOLASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ005697; CAA06682.1; -  
 CC InterPro; IPR000771; -

```
DR pfam; PF01116; F_bp_aldolase; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW Lysase; Glycolysis; Zinc.
FT METAL 83 86 ZINC (BY SIMILARITY).
FT METAL 86 86 ZINC (BY SIMILARITY).
SQ SEQUENCE 293 AA; 31401 MW; 577D2A55713B6246 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1261 LKLAKEV 1267
11111111
116 LKLAKEV 122

RESULT 41
ARGC_BACST STANDARD; PRT; 294 AA.
AC Q07906;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (EC 1.2.1.38) (AGPR) (N-
DE ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE)
DE (FRAGMENT).
GN ARGC.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 8224;
RX MEDLINE=93232760; PubMed=8473852;
RA Sakanyan V., Charlier D.R.M., Legrain C., Kochikyan A., Mett I.,
RA Pierard P., Glandsdorff N.;
RT "Primary structure, partial purification and regulation of key
RT enzymes of the acetyl cycle of arginine biosynthesis in Bacillus
RT stearothermophilus: dual function of ornithine acetyltransferase.";
RL J. Gen. Microbiol. 139:393-402(1993).
CC -!- CATALYTIC ACTIVITY: N-ACETYL-L-GLUTAMATE 5-SEMIALDEHYDE + NADP(+)
CC + ORTHOPHOSPHATE -> N-ACETYL-5-GLUTAMYL PHOSPHATE + NADPH.
CC -!- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06036; AA22196.1; -
DR InterPro; IPR000534; -
DR InterPro; IPR000706; -
DR pfam; PF01118; Semialdehyde_dh; 1.
DR PROSITE; PS01224; ARG; 1.
KW Arginine biosynthesis; Oxidoreductase; NADP.
FT NON_TER 1 1
FT ACT_SITE 98 98 BY SIMILARITY.
SQ SEQUENCE 294 AA; 32037 MW; C2A91B08313F4673 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 294;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 209 SGAGRKA 215
11111111

RESULT 42
HX7L_XENLA STANDARD; PRT; 295 AA.
ID HX7L_XENLA AC Q04281;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE HOMEOBOX PROTEIN XHOX-7.1 (FRAGMENT).
GN XHOX-7.1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91347929; PubMed=1679007;
RA Su M.-W., Suzuki H.R., Solursh M., Ramirez F.;
RT "Progressively restricted expression of a new homeobox-containing
RT gene during Xenopus laevis embryogenesis.";
RL Development 111:1179-1187(1991).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- DEVELOPMENTAL STAGE: APPEARS AT THE BEGINNING OF GASTRULATION.
CC PLATEAU BETWEEN THE NEURULA AND MIDDLE-TAILBUD STAGES, AND
CC DECREASE STEADILY THEREAFTER.
CC -!- SIMILARITY: BELONGS TO THE MSH FAMILY OF HOMEOBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58773; CAA41574.1; -
DR PIR; A60131; A60131.
DR HSP; P22808; LVND.
DR InterPro; IPR001356; -
DR pfam; PF00046; homeobox; 1.
DR PROSITE; PS00027; HOMEOBOX_1; 1.
DR PROSITE; PS00071; HOMEOBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT NON_TER 1 1
FT DNA_BIND 170 229
SQ SEQUENCE 295 AA; 32338 MW; 054331E2BC106C10 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 295;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1046 ASLYGTS 1052
11111111
261 ASLYGTS 267

RESULT 43
Y175_HELPJ STANDARD; PRT; 299 AA.
ID Y175_HELPJ AC Q9ZMQ7;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN JHP0161 PRECURSOR.
GN JHP0161.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
```

```

[1]
SEQUENCE FROM N.A.
MEDLINE-99120557; PubMed-9923682;
Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
Trust T.J.;
"Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.";
Nature 397:176-180(1999).
-1- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
STRONG, TO C.JEJUNI CBF2.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE001454; AAD05744.1;
HSSP; Q13526; IPIN.
InterPro; IPR000297;
Pfam; PF00639; Rotamase; 1.
PROSITE; PS01096; PPIC_PPIASE_1; 1.
Hypothetical protein; Isomerase; Rotamase; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 299 HYPOTHETICAL PROTEIN JHP0161.
FT DOMAIN 154 253 PPIC-LIKE.
FT SEQUENCE 299 AA; 34040 MW; 9C037BICD1110143 CRC64;
-----
Query Match 0.5%; Score 7; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 LALVGAL 23
D 8 LALVGAL 14
-----
RESULT 44
Y175_HELPY STANDARD; PRT; 299 AA.
AC P56112;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN HP0175 PRECURSOR.
GN HP0175.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
[1]
SEQUENCE FROM N.A.
STRAIN=26695 / ATCC 700392;
MEDLINE-97394467; PubMed-9252185;
Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
Fleischmann R.D., Ketchum K.A., Klenk H.-B., Gill S., Dougherty B.A.,
Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
Lofthus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
Venter J.C.;
"The complete genome sequence of the gastric pathogen Helicobacter
pylori.";
Nature 388:539-547(1997).
-1- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.

```

```

STRONG, TO C.JEJUNI CBF2.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE000538; AAD07245.1;
TIGR; HP0175;
InterPro; IPR000297;
Pfam; PF00639; Rotamase; 1.
PROSITE; PS01096; PPIC_PPIASE_1; 1.
PROSITE; PS01098; PPIC_PPIASE_2; 1.
Hypothetical protein; Isomerase; Rotamase; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 299 HYPOTHETICAL PROTEIN HP0175.
FT DOMAIN 154 253 PPIC-LIKE.
FT SEQUENCE 299 AA; 34031 MW; E65F3F2F94B11F5A CRC64;
-----
Query Match 0.5%; Score 7; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 LALVGAL 23
D 8 LALVGAL 14
-----
RESULT 45
PTB_CLOAB STANDARD; PRT; 302 AA.
ID PTB_CLOAB
AC Q05624;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE PHOSPHATE BUTYRYLTRANSFERASE (EC 2.3.1.19) (PHOSPHOTRANSBUTYRYLASE).
GN PTB.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
[1]
SEQUENCE FROM N.A.
STRAIN=NCIMB 8052;
MEDLINE-93380658; PubMed-8396545;
Oultram J.D., Burr I.D., Elmore M.J., Minton N.P.;
"Cloning and sequence analysis of the genes encoding
phosphotransbutyrylase and butyrate kinase from Clostridium
acetobutylicum NCIMB 8052.";
Gene 131:107-112(1993).
-1- FUNCTION: CATALYSES THE CONVERSION OF BUTYRYL-COA THROUGH BUTYRYL
PHOSPHATE TO BUTYRATE.
-1- CATALYTIC ACTIVITY: BUTANOYL-COA + ORTHOPHOSPHATE = COA +
BUTANOYL PHOSPHATE.
-1- PATHWAY: INVOLVED IN THE ACIDGENIC PHASE OF FERMENTATION.
-1- SIMILARITY: BELONGS TO THE PHOSPHATE ACETYLTRANSFERASE AND
BUTYRYLTRANSFERASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; L04468; AAA52080.1;
PIR; JN0794; JN0794.
InterPro; IPR002505;

```



```

DR pfam; PF01515; PTA_PTB; 1.
KW Transferase; Acyltransferase.
SQ SEQUENCE 302 AA; 32441 MW; 52C3A7BFB187C1FF CRC64;

Query Match 0.5%; Score 7; DB 1; Length 302;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 VIDGPFA 391
Db 209 VIDGPFA 215
|||||

RESULT 46
ACCA_ECOLI STANDARD; PRT; 318 AA.
ID ACCA_ECOLI
AC P30867;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ACETYL-COENZYME A CARBOXYLASE CARBOXYL TRANSFERASE SUBUNIT ALPHA
  (EC 6.4.1.2).
GN ACCA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-K12 / W3110;
RX MEDLINE-92380982; PubMed-1355089;
RA Li S.-J., Cronan J.E. Jr.;
RT "The genes encoding the two carboxyltransferase subunits of
  Escherichia coli acetyl-CoA carboxylase."
RL J. Biol. Chem. 267:16841-16847(1992).
  [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RA Yamamoto Y.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
  [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / M61655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
  Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
  Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
  Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
  [4]
RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
  Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
  Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
  Davis R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
  [5]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
  Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R.,
  Ichihara S., Mizuno T., Makino K., Nakata A., Yura T., Sampei G.,
  Mizobuchi K.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
  [6]
RP SEQUENCE OF 163-318 FROM N.A.
RC STRAIN-K12 / W3110;
RX MEDLINE-97369816; PubMed-9226257;
RA Kikuchi Y., Kojima H., Tanaka T., Takatsuka Y., Kamio Y.;
RT "Characterization of a second lysine decarboxylase isolated from
  Escherichia coli."
  [7]
J. Bacteriol. 179:4486-4492(1997);
RN [7]
RP SEQUENCE OF 1-12.
RC STRAIN-K12 / EMG2;
RX MEDLINE-97443975; PubMed-9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
  in the genome of Escherichia coli K-12."
RL Electrophoresis 18:1259-1313(1997)
CC -|- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
  CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
  CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
  TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC -|- CATALYTIC ACTIVITY: CARBOXYBIOTIN CARBOXYL CARRIER PROTEIN +
  ACETYL-COA -> BIOTIN CARBOXYL CARRIER PROTEIN + MALONYL-COA.
CC -|- PATHWAY: FIRST STEP IN LONG-CHAIN FATTY ACID SYNTHESIS.
CC -|- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
  CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
  OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
CC -|- SIMILARITY: TO THE C-TERMINUS OF MAMMALIAN PROPIONYL-COA
  CARBOXYLASE BETA CHAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
DR EMBL; M96394; AAA70370.1; -
DR EMBL; D49445; BAA08425.1; -
DR EMBL; AE000127; AAC73296.1; -
DR EMBL; U70214; AAB08614.1; -
DR EMBL; D83536; BAA77860.1; -
DR EMBL; D87518; BAA21655.1; -
DR PIR; A43452; A43452.
DR EcoGene; EG11647; acca.
DR InterPro; IPR001095; -
DR PRINTS; PR01069; ACCCTRFRASEA.
KW Fatty acid biosynthesis; Ligase.
FT INIT_MET 0
FT DOMAIN 83 118 ACYL-COA-BINDING DOMAIN (POTENTIAL).
FT CONFLICT 24 24 V -> G (IN REF. 5).
SQ SEQUENCE 318 AA; 35110 MW; 8938BE08E5D3C9AD CRC64;

Query Match 0.5%; Score 7; DB 1; Length 318;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AIVGGIA 49
Db 99 AIVGGIA 105
|||||

RESULT 47
ALDX_PIG STANDARD; PRT; 324 AA.
ID ALDX_PIG
AC P50578;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ALCOHOL DEHYDROGENASE [NADP+] (EC 1.1.1.2) (ALDEHYDE REDUCTASE).
GN AKR1A1 OR ALR OR ALR1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE-96062952; PubMed-7484379;

```

FLynn T.G., Green N.C., Bhatia M.B., El-Kabbani O.;  
 "Structure and mechanism of aldehyde reductase";  
 Adv. Exp. Med. Biol. 372:193-201(1995).  
 [2]  
 X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 MEDLINE=97469638; PubMed=9329083;  
 El-Kabbani O., Carper D.A., McGowan M.H., Devedjiev Y.,  
 Rees-Milton K.J., Flynn T.G.  
 "Studies on the inhibitor-binding site of porcine aldehyde reductase:  
 crystal structure of the holoenzyme-inhibitor ternary complex";  
 Proteins 29:186-192(1997).  
 [3]  
 X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 El-Kabbani O., Carper D.A., McGowan M.H., Ginell S.L.;  
 Submitted (JUL-1996) to the PDB data bank.  
 -1- CATALYTIC ACTIVITY: ALCOHOL + NADP(+) = ALDEHYDE + NADPH.  
 -1- SUBUNIT: MONOMER.  
 -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE FAMILY.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; U46064; AAB60266.1; -  
 PDB; 1AE4; 11-MAR-98.  
 PDB; 1CWN; 04-FEB-98.  
 InterPro: IPR001395; -  
 Pfam; PF00248; aldo\_ket\_red; 1.  
 PRINTS; PR00069; ALDKETREDTASE.  
 PROSITE; PS00062; ALDOKETO\_REDUCTASE\_2; 1.  
 PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.  
 PROSITE; PS00798; ALDOKETO\_REDUCTASE\_1; 1.  
 Oxidoreductase; NADP; Acetylation; 3D-structure.  
 INIT-MET 0 0 BY SIMILARITY.  
 MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
 ACT\_SITE 112 112 HYDROGEN-BOND DONOR (BY SIMILARITY).  
 VARIANT 164 164 S -> N  
 SEQUENCE 324 AA; 36407 MW; 4EB6080B5951E057 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 324;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 NADGTIR 412  
 | | | | |  
 Db 127 NADGTIR 133

RESULT 48  
 ID ACCA\_BACSU STANDARD; PRT; 325 AA.  
 AC 034847;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE ACETYL-COENZYME A CARBOXYLASE SUBUNIT ALPHA  
 DE (EC 6.4.1.2).  
 GN ACCA.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98048467; PubMed=9387221;  
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;  
 RT "Sequencing and functional annotation of the Bacillus subtilis genes  
 in the 200 kb rrnB-dnaB region.";

Microbiology 143:3431-3441(1997).  
 -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A  
 CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE  
 CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE  
 TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA (BY SIMILARITY).  
 -1- CATALYTIC ACTIVITY: CARBOXYBIOTIN CARBOXYL CARRIER PROTEIN +  
 ACETYL-COA = BIOTIN CARBOXYL CARRIER PROTEIN + MALONYL-COA.  
 -1- PATHWAY: FIRST STEP IN LONG-CHAIN FATTY ACID SYNTHESIS.  
 -1- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHXAMER OF BIOTIN  
 CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS  
 OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (BY SIMILARITY).  
 -1- SIMILARITY: TO THE C-TERMINUS OF MAMMALIAN PROPIONYL-COA  
 CARBOXYLASE BETA CHAIN.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; AF008220; AAC00341.1; -  
 EMBL; Z99118; CAB14880.1; -  
 Subtilist; BG12557; acca.  
 DR InterPro; IPR001095; -  
 DR PRINTS; PR01069; ACCCTRFRASEA.  
 KW Fatty acid biosynthesis; Ligase.  
 SQ SEQUENCE 325 AA; 36333 MW; 9B177DEE4A5B5864 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 325;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AIVGGIA 49  
 | | | | |  
 Db 96 AIVGGIA 102

RESULT 49  
 YGHQ\_ECOLI STANDARD; PRT; 325 AA.  
 ID YGHQ\_ECOLI  
 AC Q46841;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE HYPOTHETICAL 35.2 KDA PROTEIN IN GLCC-PIIB INTERGENIC REGION.  
 GN YGHQ.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 Science 277:1453-1474(1997).  
 RL  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; U28377; AAA69150.1; -



VARIANTS ARG-133 AND MET-275.  
 MEDLINE-91290050; PubMed-2063869;  
 RA Plawinska M., Thunell S., Holmberg L., Wetmur J.G., Desnick R.J.;  
 RT "Delta-aminolevulinic dehydratase deficient porphyria:  
 RT identification of the molecular lesions in a severely affected  
 RT homozygote.";  
 RL Am. J. Hum. Genet. 49:167-174(1991).  
 RN [6]  
 RN VARIANTS TRP-240 AND THR-274.  
 RX MEDLINE-922325256; PubMed-1569184;  
 RA Ishida N., Fujita H., Fukuda Y., Noguchi T., Doss M., Kappas A.,  
 RA Sassa S.;  
 RT "Cloning and expression of the defective genes from a patient with  
 RT delta-aminolevulinic dehydratase porphyria.";  
 RL J. Clin. Invest. 89:1431-1437(1992).  
 CC -!- CATALYTIC ACTIVITY: 2 5-AMINOLEVULINATE - PORPHOBILINOGEN +  
 CC 2 H(2)O.  
 CC -!- COFACTOR: ZINC.  
 CC -!- PATHWAY: SECOND STEP IN PORPHYRIN AND HEME BIOSYNTHESIS.  
 CC -!- SUBUNIT: HOMOOCTAMER.  
 CC -!- POLYMORPHISM: THERE ARE TWO COMMON ALLELES OF ALAD. INDIVIDUALS  
 CC HETEROZYGOUS OR HOMOZYGOUS FOR THE 2ND ALLELE HAVE SIGNIFICANTLY  
 CC HIGHER BLOOD LEAD LEVELS THAN DO 1ST ALLELE HOMOZYGOES WHEN  
 CC EXPOSED TO ENVIRONMENTAL LEAD.  
 CC -!- DISEASE: DEFECTS IN ALAD ARE THE CAUSE OF ACUTE HEPATIC PORPHYRIA.  
 CC -!- SIMILARITY: BELONGS TO THE ALADH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M13928; AAA51687.1; -;  
 DR EMBL; X64467; CAA45796.1; -;  
 DR EMBL; S99468; AAC60581.1; -;  
 DR EMBL; S99471; AAC60582.1; -;  
 DR PIR; A26478; A26478.  
 DR SWISS-2DPAGE; P13716; HUMAN.  
 DR MIM; 125270; -;  
 DR InterPro; IPR001731; -;  
 DR Pfam; PF00490; ALAD; 1.  
 DR PRINTS; PR00144; DALDHYDRASE.  
 DR PROSITE; PS00169; D\_ALA\_DEHYDRATASE; 1.  
 KW Porphyrin biosynthesis; Heme biosynthesis; Lyase; Zinc;  
 KW Disease mutation; Polymorphism.  
 FT DOMAIN 119 137  
 FT ACT\_SITE 252 252  
 FT VARIANT 59 59 K -> N (IN ALLELE 2; 10% OF POPULATION).  
 FT VARIANT 133 133 /FTIQ-VAR\_003633.  
 FT VARIANT 240 240 /FTIQ-VAR\_003634.  
 FT VARIANT 274 274 /FTIQ-VAR\_003635.  
 FT VARIANT 275 275 /FTIQ-VAR\_003636.  
 FT VARIANT 275 275 /FTIQ-VAR\_003637.  
 FT SEQUENCE 330 AA; 36295 MW; E005F305F6D9403 CRC64;  
 -----  
 Query Match 0.5%; Score 7; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 948 TATTTLN 954  
 DB 21 TATTTLN 27  
 RESULT 52

ARGC\_ECOLI STANDARD; PRT; 334 AA.  
 ID ARGC\_ECOLI  
 AC P11446;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (EC 1.2.1.38) (AGPR) (N-  
 DE ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE).  
 GN ARGC.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE-89121510; PubMed-2851495;  
 RA Parsot C., Boyen A., Cohen G.N., Glansdorff N.;  
 RT "Nucleotide sequence of Escherichia coli argB and argC genes:  
 RT comparison of N-acetylglutamate kinase and  
 RT N-acetylglutamate-gamma-semialdehyde dehydrogenase with homologous  
 RT and analogous enzymes.";  
 RL Gene 68:275-283(1988).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RC MEDLINE-94089392; PubMed-8265357;  
 RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,  
 RA Daniels D.L.;  
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the  
 RT region from 89.2 to 92.8 minutes.";  
 RL Nucleic Acids Res. 21:5408-5417(1993).  
 RN [3]  
 RN SEQUENCE OF 1-48 FROM N.A.  
 RC STRAIN-K12;  
 RC MEDLINE-83143275; PubMed-6761650;  
 RA Piette J., Cunin R., Boyen A., Charlier D.R.M., Crabeel M.,  
 RA van Vliet F., Glansdorff N., Squires C., Squires C.L.;  
 RT "The regulatory region of the divergent argECB operon in Escherichia  
 RT coli K-12.";  
 RL Nucleic Acids Res. 10:8031-8048(1982).  
 RN [4]  
 RN SEQUENCE OF 1-19 FROM N.A.  
 RC STRAIN-K12;  
 RC MEDLINE-92202162; PubMed-1551850;  
 RA Meinel T., Schmitt E., Mechulam Y., Blanquet S.;  
 RT "Structural and biochemical characterization of the Escherichia coli  
 RT argE gene product.";  
 RL J. Bacteriol. 174:2323-2331(1992).  
 CC -!- CATALYTIC ACTIVITY: N-ACETYL-L-GLUTAMATE 5-SEMIALDEHYDE + NADP(+) +  
 CC + ORTHOPHOSPHATE -> N-ACETYL-5-GLUTAMYL PHOSPHATE + NADPH.  
 CC -!- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.  
 CC -!- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M21446; AAA23477.1; -;  
 DR EMBL; J01587; AAB59146.1; -;  
 DR EMBL; X55417; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U00006; AAC43064.1; -;  
 DR EMBL; AE000470; AAC76940.1; -;  
 DR FIR; J03032; RDECEP.  
 DR Ecogene; EGI0065; argC.  
 DR InterPro; IPR000534; -;  
 DR InterPro; IPR000706; -;  
 DR Pfam; PF01118; Semialdehyde\_dh; 1.  
 DR PROSITE; PS01224; ARGC; 1.  
 KW Arginine biosynthesis; Oxidoreductase; NADP.

```
FT ACT_SITE 154 154 BY SIMILARITY.
SQ SEQUENCE 334 AA; 35952 MW; 67AC195ECE1C4789 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKA 215
    |||||
DB 187 SGAGRKA 193

RESULT 54
HRCA_MYCCA
ID HRCA_MYCCA STANDARD; PRT; 334 AA.
AC P71498;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HEAT-INDUCIBLE TRANSCRIPTION REPRESSOR HRCA.
GN HRCA.
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Entomoplasmataceae.
OX NCBI_TaxID=2095;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 25416;
RC MEDLINE=971148974; PubMed=8995799;
RA Falah M., Gupta R.S.;
RT "Phylogenetic analysis of mycoplasmas based on Hsp70 sequences:
cloning of the dnaK (hsp70) gene region of Mycoplasma capricolum.";
RL Int. J. Syst. Bacteriol. 47:38-45(1997).
CC -!- FUNCTION: NEGATIVE REGULATOR OF CLASS I HEAT SHOCK GENES (GRPE-
DNAK-DNAJ AND GROELS OPERONS). PREVENTS HEAT-SHOCK INDUCTION OF
THESE OPERONS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HRCA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U51235; AAB09428.1; -
DR InterPro; IPR002571; -
DR Pfam; PF01628; HRCA; 1.
KW Transcription regulation; Repressor; Heat shock.
SQ SEQUENCE 334 AA; 38637 MW; 2C44C27A39ADB922 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 SSTVLTL 222
    |||||
DB 306 SSTVLTL 312

RESULT 54
YEHA_ECOLI
ID YEHA_ECOLI STANDARD; PRT; 344 AA.
AC P33340;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 36.9 KDA PROTEIN IN GATY-MRP INTERGENIC REGION PRECURSOR.
GN YEHA.
OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / BHB2600;
RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
Church G.M.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aliba H., Baba T., Fujita K., Hayashi K., Inada T.,
Isono K., Kasai H., Kimura S., Kitagawa M., Kitagawa M.,
Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
Salto N., Sampei G., Seki Y., Sivasubram S., Tagami H.,
Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00007; AAA60472.1; -
DR EMBL; AE000300; AAC75169.1; -
DR EMBL; D90848; BAAL5974.1; -
DR Ecogene; Ecol1987; yeHA.
KW Hypothetical protein; Signal.
FT CHAIN 1 20 POTENTIAL.
FT SIGNAL 1 20 HYPOTHETICAL PROTEIN YEHA.
SQ SEQUENCE 344 AA; 36885 MW; CFC3A72B0227A033 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 861 NALAQNA 867
    |||||
DB 172 NALAQNA 178

RESULT 55
RIR2_TREPA
ID RIR2_TREPA STANDARD; PRT; 351 AA.
AC O83092;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE BETA CHAIN (EC 1.17.4.1)
DE (RIBONUCLEOTIDE REDUCTASE).
GN NRDB OR TP0053.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Attiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -1- FUNCTION: CATALYZES THE BIOSYNTHESIS OF DEOXYRIBONUCLEOTIDES FROM
CC THE CORRESPONDING RIBONUCLEOTIDES, PRECURSORS THAT ARE NECESSARY
CC FOR DNA SYNTHESIS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2'-DEOXYRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED
CC THIOREDOXIN + H(2)O = RIBONUCLEOSIDE DIPHOSPHATE + REDUCED
CC THIOREDOXIN.
CC -1- COFACTOR: CONTAINS TWO IRON IONS (BY SIMILARITY).
CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001190; AAC65049.1;
DR TIGR: TP0053;
DR InterPro: IPR000358;
DR Pfam: PF00268; ribonuc_red; 1.
DR PROSITE: PS00368; RIBED_SMALL; 1.
KW Oxidoreductase; DNA replication; Iron.
FT METAL 94 94 IRON 1 (BY SIMILARITY).
FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 127 127 IRON 1 (BY SIMILARITY).
FT METAL 191 191 IRON 2 (BY SIMILARITY).
FT METAL 225 225 IRON 2 (BY SIMILARITY).
FT METAL 228 228 IRON 2 (BY SIMILARITY).
FT ACT_SITE 131 131 BY SIMILARITY.
SQ SEQUENCE 351 AA; 41198 MW; 5B373A7FCBBF049A CRC64;

Query Match 0.5%; Score 7; DB 1; Length 351;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 842 GGNTNL 848
DB 29 GGNTNL 35
|||||

RESULT 56
MRAY_HELPY STANDARD; PRT; 353 AA.
AC O25235;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE PHOSPHO-N-ACETYLURAMONYL-PENTAPEPTIDE-TRANSFERASE (EC 2.7.8.13)
DE (UDP-MURNAc-PENTAPEPTIDE PHOSPHOTRANSFERASE).
GN MRAY OR HP0493.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.

```

```

OX RN NCBI_TaxID=210;
RN RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Loftus K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Nelson B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA Kennedy K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE
CC BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: UDP-N-ACETYLMURAMYL-L-ALANYL-D-GLUTAMYL-L-
CC LYSYL-D-ALANYL-D-ALANINE + UNDECAPRENYL PHOSPHATE = UMP +
CC N-ACETYLMURAMYL-L-ALANYL-D-GLUTAMYL-L-LYSYL-D-ALANYL-D-ALANINE-
CC DIPHOSPHOUNDECAPRENOL.
CC -1- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. MRAY
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000564; AAD07559.1;
DR TIGR: HP0493;
DR InterPro: IPR000715;
DR Pfam: PF00953; Glycos_transf_4; 1.
DR PROSITE: PS01347; MRAY_1; 1.
DR PROSITE: PS01348; MRAY_2; 1.
KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane.
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 88 108 POTENTIAL.
FT TRANSMEM 129 149 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 193 213 POTENTIAL.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
FT TRANSMEM 332 352 POTENTIAL.
SQ SEQUENCE 353 AA; 39169 MW; 1BF7B35A9F29A91A CRC64;

Query Match 0.5%; Score 7; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VSLALVG 21
DB 235 VSLALVG 241
|||||

RESULT 57
YOV4_CAEEL STANDARD; PRT; 367 AA.
AC O22618;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 40.8 KDA PROTEIN T20H4.4 IN CHROMOSOME III.

```

GN T20H4.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Du 2.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE ADENOSINE DEAMINASE EDITASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U00037; AAA50661.1; -  
 DR WormPep; T20H4.4; CE00831.  
 DR PROSITE; PS0141; A\_DEAMIN\_EDITASE; 1.  
 KW Hypothetical protein; Hydrolase; Zinc.  
 FT METAL 63 63 ZINC (BY SIMILARITY).  
 FT ACT\_SITE 65 65 BY SIMILARITY.  
 FT METAL 116 116 ZINC (BY SIMILARITY).  
 FT METAL 177 177 ZINC (BY SIMILARITY).  
 SQ SEQUENCE 367 AA; 40828 MW; DA7722062324589 CRC64;  
 -----  
 Query Match 0.5%; Score 7; DB 1; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 QY 1158 NALVLKP 1164  
 DB 97 NALVLKP 103  
 -----  
 RESULT 58  
 YXER\_BACSU  
 ID YXER\_BACSU STANDARD; PRT; 370 AA.  
 AC P54957;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHETICAL 38.4 KDA PROTEIN IN IDH-DEOR INTERGENIC REGION.  
 GN YXER OR LP6A.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / BGSC1A1;  
 RX MEDLINE=97021444; PubMed=8867804;  
 RA Yoshida K.-I., Fujimura M., Yanai N., Fujita Y.;  
 RT Cloning and sequencing of a 23-kb region of the Bacillus subtilis  
 RT genome between the *iol* and *hut* operons.;  
 RL Genes Res. 2:295-301(1995).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -!- SIMILARITY: TO S.TYPHIMURIUM ETHANOLAMINE UTILIZATION PROTEIN  
 CC EUTH.  
 CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; D45912; BAA08334.1; -  
 DR EMBL; Z99124; CAB15981.1; -  
 DR Subtilist; BG11894; yxer.  
 KW Hypothetical protein; Transmembrane; Transport.  
 FT TRANSMEM 6 26 POTENTIAL.  
 FT TRANSMEM 49 69 POTENTIAL.  
 FT TRANSMEM 79 99 POTENTIAL.  
 FT TRANSMEM 111 131 POTENTIAL.  
 FT TRANSMEM 143 163 POTENTIAL.  
 FT TRANSMEM 167 187 POTENTIAL.  
 FT TRANSMEM 206 226 POTENTIAL.  
 FT TRANSMEM 236 256 POTENTIAL.  
 FT TRANSMEM 307 327 POTENTIAL.  
 FT TRANSMEM 333 353 POTENTIAL.  
 SQ SEQUENCE 370 AA; 38414 MW; E7BF1585EB600DF8 CRC64;  
 -----  
 Query Match 0.5%; Score 7; DB 1; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 83;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 QY 467 GGYALAG 473  
 DB 90 GGYALAG 96  
 -----  
 RESULT 59  
 OPFC\_MYCPN  
 ID OPFC\_MYCPN STANDARD; PRT; 376 AA.  
 AC P75553;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE OLIGOPEPTIDE TRANSPORT SYSTEM PERMEASE PROTEIN OPFC.  
 GN OPFC OR MPN216 OR MP615.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR OLIGOPEPTIDES; PROBABLY RESPONSIBLE FOR THE TRANSLOCATION OF  
 CC THE SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -!- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-  
 CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE OPFC  
 CC SUBFAMILY.  
 CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF000058; AAB96263.1; -  
 DR InterPro; IPR000515; -  
 DR Pfam; PF00528; BPD\_transp; 1.  
 DR PROSITE; PS00402; BPD\_TRANSP\_INN\_MEMBR; 1.  
 KW Transport; Peptide transport; Transmembrane.  
 FT TRANSMEM 46 86 POTENTIAL.  
 FT TRANSMEM 149 169 POTENTIAL.  
 FT TRANSMEM 173 193 POTENTIAL.  
 FT TRANSMEM 209 229 POTENTIAL.

```

FT TRANSMEM 242 262 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 341 361 POTENTIAL.
SQ SEQUENCE 376 AA; 41233 MW; FD3042160B7F9797 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 860 NNALAQN 866
Db 78 NNALAQN 84

RESULT 60
Y823_MYCTU
ID Y823_MYCTU STANDARD; PRT; 389 AA.
AC O53835;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE HYPOTHETICAL 41.4 KDA PROTEIN RV0823C.
GN RV0823C OR MTV043.i5C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV.
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felkwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544(1998).
CC -!- SIMILARITY: BELONGS TO THE UPF0034 (NIPR3/SMM1) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL022004; CAAL17629.1;
CC TubercuList; RV0823C;
CC InterPro; IPR001269;
CC Pfam; PF01207; UPF0034; 1.
CC PROSITE; PS01136; UPF0034; 1.
CC Hypothetical protein.
SQ SEQUENCE 389 AA; 41388 MW; 11F5EB06E561E0DB CRC64;

Query Match 0.5%; Score 7; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 VGTVSGL 60
Db 49 VGTVSGL 55

RESULT 61
NUSA_HELPY

```

```

ID NUSA_HELPY STANDARD; PRT; 395 AA.
AC Q92JAG;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE N UTILIZATION SUBSTANCE PROTEIN A HOMOLOG.
GN NUSA OR HP1407.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- FUNCTION: COULD PARTICIPATES IN BOTH THE TERMINATION AND
CC ANTITERMINATION OF TRANSCRIPTION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE NUSA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001562; AAD06980.1;
CC InterPro; IPR000958;
CC InterPro; IPR003029;
CC Pfam; PF00013; KH-domain; 2.
KW Transcription termination; Repeat.
FT DOMAIN 137 201
FT DOMAIN 243 291
FT DOMAIN 331 378
FT SEQUENCE 395 AA; 44737 MW; B7536A8674BDD3D7 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 395;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 429 IGKGGVN 435
Db 343 IGKGGVN 349

RESULT 62
NUSA_HELPY
ID NUSA_HELPY STANDARD; PRT; 395 AA.
AC P55977;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE N UTILIZATION SUBSTANCE PROTEIN A HOMOLOG.
GN NUSA OR HP1514.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,

```



RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodak A.,  
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
 RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,  
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*  
 RT *pylori*.";  
 RL Nature 388:539-547(1997).  
 CC -!- FUNCTION: COULD PARTICIPATES IN BOTH THE TERMINATION AND  
 CC ANTI-TERMINATION OF TRANSCRIPTION (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE NUSA FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 'SL MOTIF' DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 2 KH DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE000649; AAD08555.1;  
 DR TIGR: HPI514;  
 KW Transcription termination; Repeat.  
 FT DOMAIN 137 201 SL.  
 FT DOMAIN 243 291 KH.  
 FT DOMAIN 331 378 KH.  
 SQ SEQUENCE 395 AA; 44649 MW; DDFB53B2FC53E2B CRC64;

Query Match 0.5%; Score 7; DB 1; Length 395;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 IGRGGVN 435  
 Db 343 IGRGGVN 349

## RESULT 63

ID CBPB\_PIG STANDARD; PRT; 401 AA.  
 AC F09955;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CARBOXYPEPTIDASE B PRECURSOR (EC 3.4.17.2).  
 GN CPB.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE OF 1-103.  
 RX MEDLINE-91208150; PubMed-2018774;  
 RA Burgos F.J., Salva M., Vallegas V., Soriano F., Mendez E.,  
 RA Aviles F.X.;  
 RT "Analysis of the activation process of porcine procarboxypeptidase B  
 RT and determination of the sequence of its activation segment.";  
 RL Biochemistry 30:4082-4089(1991).  
 RN [2]  
 RP PRELIMINARY SEQUENCE OF 1-38.  
 RX MEDLINE-85279427; PubMed-4026847;  
 RA Aviles F.X., Vendrell J., Burgos F.J., Soriano F., Mendez E.;  
 RT "Sequential homologies between procarboxypeptidases A and B from  
 RT porcine pancreas.";  
 RL Biochem. Biophys. Res. Commun. 130:97-103(1985).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE-91114690; PubMed-1989878;

RA Coll M., Guasch A., Aviles F.X., Huber R.;  
 RT "Three-dimensional structure of porcine procarboxypeptidase B: a  
 RT structural basis of its inactivity.";  
 RL EMBO J. 10:1-9(1991).  
 RN [4]  
 RP STRUCTURE BY NMR OF ACTIVATION PEPTIDE, AND SEQUENCE OF 1-81.  
 RX MEDLINE-91027767; PubMed-2223783;  
 RA Vendrell J., Wider G., Aviles F.X., Wuethrich K.;  
 RT "Sequence-specific 1H NMR assignments and determination of the  
 RT secondary structure for the activation domain isolated from  
 RT pancreatic procarboxypeptidase B.";  
 RL Biochemistry 29:7515-7522(1990).  
 RN [5]  
 RP STRUCTURE BY NMR OF ACTIVATION PEPTIDE.  
 RX MEDLINE-91114693; PubMed-1989879;  
 RA Vendrell J., Billeter M., Wider G., Aviles F.X., Wuethrich K.;  
 RT "The NMR structure of the activation domain isolated from porcine  
 RT procarboxypeptidase B.";  
 RL EMBO J. 10:11-15(1991).  
 RN [6]  
 RP STRUCTURE BY NMR OF ACTIVATION PEPTIDE.  
 RX MEDLINE-93044373; PubMed-1422143;  
 RA Billeter M., Vendrell J., Wider G., Aviles F.X., Coll M., Guasch A.,  
 RA Huber R., Wuethrich K.;  
 RT "Comparison of the NMR solution structure with the X-ray crystal  
 RT structure of the activation domain from procarboxypeptidase B.";  
 RL J. Biomol. NMR 2:1-10(1992).  
 CC -!- CATALYTIC ACTIVITY: PEPTIDYL-L-LYSINE/ARGININE + H(2)O = PEPTIDE +  
 CC L-LYSINE/ARGININE.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M14; ALSO KNOWN AS THE  
 CC ZINC CARBOXYPEPTIDASE FAMILY.  
 CC -!- DATABASE: NAME=Worthington enzyme manual;  
 CC WWW="http://www.worthington-biochem.com/manual/C/COB.html".  
 DR PIR: B29181; B29181.  
 DR PDB: 1PBA; 31-OCT-93.  
 DR PDB: 1NSA; 24-DEC-97.  
 DR MEROPS: M14\_003;  
 DR InterPro: IPR000834;  
 DR Pfam: PF00246; Zn\_carboxypept; 1.  
 DR PRINTS: PR00765; CRBOXYPTASEA.  
 DR PROSITE: PS00132; CARBOXYPEPT\_ZN\_1; 1.  
 DR PROSITE: PS00133; CARBOXYPEPT\_ZN\_2; 1.  
 DR Hydrolase: Carboxypeptidase; Metalloprotease; Zinc; Zymogen;  
 KW 3D-structure.  
 FT PROPEP 1 95 ACTIVATION PEPTIDE.  
 FT CHAIN 96 401 CARBOXYPEPTIDASE B.  
 FT DISULFID 158 171  
 FT DISULFID 230 253  
 FT DISULFID 244 258  
 FT METAL 161 161 ZINC.  
 FT METAL 164 164 ZINC.  
 FT METAL 289 289 NUCLEOPHILE.  
 FT ACT\_SITE 341 341  
 FT ACT\_SITE 363 363  
 FT STRAND 14 17  
 FT HELIX 20 31  
 FT TURN 32 32  
 FT HELIX 43 45  
 FT STRAND 50 52  
 FT HELIX 61 69  
 FT TURN 70 72  
 FT STRAND 75 76  
 SQ SEQUENCE 401 AA; 45713 MW; 53129AF159A26348 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1046 ASLYGTS 1052  
 Db 324 ASLYGTS 330

```

RESULT 64
COAT_BOOLV          STANDARD;          PRT;          403 AA.
ID COAT_BOOLV
AC P12869;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE COAT PROTEIN PRECURSOR.
GN ALPHA.
OS Boollarra virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nodaviridae;
OC Alphadodavirus.
OX NCBI_TaxID=12286;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90016821; PubMed=27981110;
RA Dasgupta R., Sgro J.-Y.;
RT "Nucleotide sequences of three Nodavirus RNA2's: the messengers for
their coat protein precursors."
RL Nucleic Acids Res. 17:7525-7526(1989).
RN [2]
RP SIMILARITY TO OTHER NODAVIRUSES.
RX MEDLINE=90339486; PubMed=2116525;
RA Kaesberg P., Dasgupta R., Sgro J.-Y., Wery J.P., Selling B.H.,
RA Husur M.V., Johnson J.E.;
RT "Structural homology among four nodaviruses as deduced by sequencing
and X-ray crystallography."
RL J. Mol. Biol. 214:423-435(1990).
CC -1- PTM: ENZYMATICALLY CLEAVED INTO COAT PROTEINS BETA AND GAMMA.
CC HOWEVER, THE EXACT CLEAVAGE SITE HAS NOT BEEN DETERMINED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A6.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X15960; CAA34082.1; -.
DR PIR; A34011; VCBBL.
DR PIR; S11038; S11038.
DR HSSP; F04329; 2BBV.
DR MEROPS; A06.001; -.
DR InterPro; IPR000696; -.
DR Pfam; PF01829; Peptidase A6; 1.
DR PRINTS; PR00863; NODAVIRPTASE.
KW Coat protein; Hydrolase; Aspartyl protease.
SQ SEQUENCE 403 AA; 43357 MW; 73C3533D238B1EE9 CRC64;

Query Match          0.5%; Score 7; DB 1; Length 403;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 VGGIATG 51
   |||||
Db 384 VGGIATG 390

RESULT 65
KYK2_DICDI
ID KYK2_DICDI          STANDARD;          PRT;          410 AA.
AC P18161;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE TYROSINE-PROTEIN KINASE 2 (EC 2.7.1.112) (FRAGMENT).
GN PKB OR DPYK2.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.

```

```

OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
Dictyostelium discoideum."
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
PROTEIN TYROSINE PHOSPHATE.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES BUT ALSO TO
SERINE/THREONINE PROTEIN KINASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; M33784; AAA33203.1; -.
DR PIR; B35670; B35670.
DR HSSP; P11362; IFGI.
DR DictyDb; DD03011; pykb.
DR InterPro; IPR000719; -.
DR InterPro; IPR001245; -.
DR Pfam; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT NON_TER
FT DOMAIN 108 381 PROTEIN KINASE.
FT NF_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E93918B605B9AEC1 CRC64;

Query Match          0.5%; Score 7; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 EEQFKER 766
   |||||
Db 143 EEQFKER 149

RESULT 66
PRIL_CAEEL
ID PRIL_CAEEL          STANDARD;          PRT;          410 AA.
AC P34471;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROBABLE DNA PRIMASE SMALL SUBUNIT (EC 2.7.7.-).
GN F58A4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,

```

RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
CC -1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL  
CC RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS  
CC DNA REPLICATION (BY SIMILARITY).  
CC -1- SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.  
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE SMALL SUBUNIT  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J22179; CAA80161.1; -;  
CC PIR; S40976; S40976.  
CC WormPep; F58A4.4; CE00220.  
CC InterPro; IPR002755; -;  
CC Pfam; PF01896; DNA\_primase\_S; 1.  
CC Transferrase; DNA replication; DNA-directed RNA polymerase; Primosome.  
CC ACT\_SITE 43  
CC ACT\_SITE 106 106 POTENTIAL.  
CC ACT\_SITE 108 108 POTENTIAL.  
CC METAL 118 118 POTENTIAL.  
CC METAL 119 119 POTENTIAL.  
CC METAL 125 125 POTENTIAL.  
CC METAL 128 128 POTENTIAL.  
CC SEQUENCE 410 AA; 48097 MW; 833C4EE597D527D6 CRC64;  
CC -----

Query Match 0.5%; Score 7; DB 1; Length 410;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 746 NVSTGNT 752  
| | | | |  
Db 317 NVSTGNT 323

RESULT 67  
ACRO\_PIG STANDARD; PRT; 415 AA.  
AC P08001; P08000;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE ACROSIN PRECURSOR (EC 3.4.21.10) (53 KDA FUCOSE-BINDING PROTEIN).  
GN ACR.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN SEQUENCE FROM N.A.  
RX MEDLINE=89308595; PubMed=2745422;  
RA Baba T., Kashiwabara S.I., Watanabe K., Itoh H., Michikawa Y.,  
RA Kimura K., Takada M., Fukamizu A., Arai Y.;  
RT "Activation and maturation mechanisms of boar acrosin zymogen based  
RT on the deduced primary structure.";  
RL J. Biol. Chem. 264:11920-11927(1989).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP TISSUE-Testis;  
RX MEDLINE=89325301; PubMed=2502391;  
RA Adham I.M., Maier W.-M., Hoyer-Fender S., Tsaousidou S., Engel W.,

RA Klemm U.;  
RT "Molecular cloning of preproacrosin and analysis of its expression  
RT pattern in spermatogenesis.";  
RL Eur. J. Biochem. 182:563-568(1989).  
RN [3]  
RN SEQUENCE OF 17-39.  
RC TISSUE-Sperm;  
RX MEDLINE=84261484; PubMed=6378631;  
RA Fock-Nuezel R., Lottspeich F., Henschen A., Mueller-Esterl W.;  
RT "Boar acrosin is a two-chain molecule. Isolation and primary  
RT structure of the light chain; homology with the pro-part of other  
RT serine proteinases.";  
RL Eur. J. Biochem. 141:441-446(1984).  
RN [4]  
RN SEQUENCE OF 25-91.  
RC TISSUE-Sperm;  
RX MEDLINE=81115822; PubMed=7007202;  
RA Fock-Nuezel R., Lottspeich F., Henschen A., Mueller-Esterl W.,  
RA Fritz H.;  
RT "N-Terminal amino acid sequence of boar sperm acrosin. Homology with  
RT other serine proteinases.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 361:1823-1828(1980).  
RN [5]  
RN SEQUENCE OF 17-32 AND 40-53.  
RC TISSUE-Sperm;  
RX MEDLINE=88083633; PubMed=3480243;  
RA Toepfer-Petersen E., Henschen A.;  
RT "Acrosin shows zona and fucose binding, novel properties for a serine  
RT proteinase.";  
RL FEBS Lett. 226:38-42(1987).  
RN [6]  
RN SEQUENCE OF 17-40.  
RX MEDLINE=90253655; PubMed=2111146;  
RA Cechova D., Toepfer-Petersen E., Zucker A., Jonakova V.;  
RT "Is spermminogen a modified proacrosin? Isolation, purification, and  
RT partial characterization of low-molecular-mass boar proacrosin.";  
RL Biol. Chem. Hoppe-Seyler 371:317-323(1990).  
RN [7]  
RN DISULFIDE BONDS, CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.  
RX MEDLINE=91085546; PubMed=2261983;  
RA Toepfer-Petersen E., Calvete J.J., Schaefer W., Henschen A.;  
RT "Complete localization of the disulfide bridges and glycosylation  
RT sites in boar sperm acrosin.";  
RL FEBS Lett. 275:139-142(1990).  
CC -1- FUNCTION: ACROSIN IS THE MAJOR PROTEASE OF MAMMALIAN SPERMATOZOA.  
CC IT IS A SERINE PROTEASE OF TRYPSIN-LIKE CLEAVAGE SPECIFICITY. IT  
CC IS SYNTHESIZED IN A ZYMOGEN FORM, PROACROSIN AND STORED IN THE  
CC ACROSOME.  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF ARG- AND LYS-BONDS; PREFERENTIAL  
CC CLEAVAGE ARG-XAA >> LYS-LYS >> LYS-XAA.  
CC -1- SUBUNIT: HEAVY CHAIN (CATALYTIC) AND A LIGHT CHAIN LINKED BY TWO  
CC DISULFIDE BONDS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J04950; AAA31131.1; -;  
CC EMBL; X14844; CAA32948.1; -;  
CC PIR; A34170; A34170.  
CC PIR; S02428; S02428.  
CC PIR; S04940; S04940.  
CC PIR; S08994; S08994.  
CC MEROPS; S01.223; -;  
CC InterPro; IPR001254; -;  
CC InterPro; IPR001314; -;

DR pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Sperm; Signal.  
 FT SIGNAL 1 16  
 FT CHAIN 17 415 ACROSIN.  
 FT CHAIN 17 39 ACROSIN LIGHT CHAIN.  
 FT CHAIN 40 338 ACROSIN HEAVY CHAIN.  
 FT PROPEP 339 415 PRO-RICH.  
 FT DISULFID 22 152 INTERCHAIN.  
 FT DISULFID 26 160 INTERCHAIN.  
 FT DISULFID 71 87  
 FT DISULFID 175 244  
 FT DISULFID 207 223  
 FT DISULFID 234 264  
 FT CARBOHYD 19 19  
 FT CARBOHYD 208 208  
 FT ACT\_SITE 86 86  
 FT ACT\_SITE 140 140  
 FT ACT\_SITE 238 238  
 FT CONFLICT 8 8  
 FT CONFLICT 211 211  
 FT CONFLICT 217 218  
 FT CONFLICT 347 347  
 FT CONFLICT 389 389  
 FT CONFLICT 399 402  
 FT CONFLICT 403 415  
 FT CONFLICT 415 AA; 45387 MW; 5AD11900B7E95772 CRC64;  
 SQ SEQUENCE 415 AA; 45387 MW; 5AD11900B7E95772 CRC64;  
 Query Match 0.5%; Score 7; DB 1; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 92;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 994 SFAKRLQ 1000  
 Db 378 SFAKRLQ 384  
 RESULT 68  
 ACRO\_HUMAN STANDARD; PRT; 421 AA.  
 AC P10323;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE ACROSIN PRECURSOR (EC 3.4.21.10).  
 GN ACR OR ACRS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=89153568; PubMed=2493394;  
 RA Baba T., Watanabe K., Kashiwabara S.-I., Arai Y.;  
 RT "Primary structure of human proacrosin deduced from its cDNA  
 sequence.";  
 RL FEBS Lett. 244:296-300(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=90306003; PubMed=2114285;  
 RA Keime S., Adham I.M., Engel W.;  
 RT "Nucleotide sequence and exon-intron organization of the human  
 proacrosin gene.";  
 RL Eur. J. Biochem. 190:195-200(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92331639; PubMed=1628652;  
 RA Vazquez-Levin M.H., Reventos J., Gordon J.W.;

RT "Molecular cloning, sequencing and restriction mapping of the genomic  
 sequence encoding human proacrosin.";  
 RL Eur. J. Biochem. 207:23-26(1992).  
 RN [4]  
 RA DISCUSSION ON ABOVE PAPER.  
 RP Adham I.A., Spitzer U., Schloesser M., Kremling H., Keime S.,  
 RA Engel W.;  
 RL Eur. J. Biochem. 207:27-28(1992).  
 CC -1- FUNCTION: ACROSIN IS THE MAJOR PROTEASE OF MAMMALIAN SPERMATOZOA.  
 CC IT IS A SERINE PROTEASE OF TRYPSIN-LIKE CLEAVAGE SPECIFICITY, IT  
 CC IS SYNTHESIZED IN A ZYMOGEN FORM, PROACROSIN AND STORED IN THE  
 CC ACROSOME.  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF ARG- AND LYS-BONDS; PREFERENTIAL  
 CC CLEAVAGE ARG-XAA >> LYS-LYS >> LYS-XAA.  
 CC -1- SUBUNIT: HEAVY CHAIN (CATALYTIC) AND A LIGHT CHAIN LINKED BY TWO  
 CC DISULFIDE BONDS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Y00970; CAA68784.1; -  
 DR EMBL; X54017; CAA37964.1; -  
 DR EMBL; X54018; CAA37964.1; JOINED.  
 DR EMBL; X54019; CAA37964.1; JOINED.  
 DR EMBL; X54020; CAA37964.1; JOINED.  
 DR EMBL; M77378; AAA51572.1; -  
 DR EMBL; M77379; AAA51573.1; -  
 DR EMBL; M77380; AAA51574.1; -  
 DR EMBL; M77381; AAA51575.1; -  
 DR EMBL; X66188; CAA46956.1; -  
 DR EMBL; X54018; CAA46956.1; JOINED.  
 DR EMBL; X54019; CAA46956.1; JOINED.  
 DR EMBL; X54020; CAA46956.1; JOINED.  
 DR PIR; S03330; S03330.  
 DR PIR; S11674; S11674.  
 DR PIR; S12063; S12063.  
 DR PIR; S23499; S23499.  
 DR MEROPS; S01.223; -  
 DR MIM; 102480; -  
 DR InterPro; IPR001254; -  
 DR InterPro; IPR001314; -  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Sperm; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 421 ACROSIN.  
 FT CHAIN 20 42 ACROSIN LIGHT CHAIN.  
 FT CHAIN 43 ? ACROSIN HEAVY CHAIN.  
 FT PROPEP ? 421 PRO-RICH.  
 FT DISULFID 25 154 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 29 162 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 73 89 BY SIMILARITY.  
 FT DISULFID 177 246 BY SIMILARITY.  
 FT DISULFID 209 225 BY SIMILARITY.  
 FT DISULFID 236 266 BY SIMILARITY.  
 FT CARBOHYD 22 22 N-LINKED (GLCNAC... ) (BY SIMILARITY).  
 FT CARBOHYD 210 210 N-LINKED (GLCNAC... ) (BY SIMILARITY).  
 FT ACT\_SITE 88 88 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 142 142 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 240 240 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CONFLICT 64 64 R -> T (IN REF. 1).  
 FT CONFLICT 120 120 L -> V (IN REF. 1).  
 FT CONFLICT 166 166 F -> L (IN REF. 1).  
 FT CONFLICT 268 268 L -> R (IN REF. 1).

FT CONFLICT 345 345 P -> R (IN REF. 1).  
 SQ SEQUENCE 421 AA; 45799 MW; 62EB47DC25B4FB5D CRC64;

Query Match 0.5%; Score 7; DB 1; Length 421;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 994 SFAKRLQ 1000  
 DB 383 SFAKRLQ 389

RESULT 69  
 ODP2\_STAUB STANDARD; PRT; 430 AA.  
 AC Q59821;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE DIHYDROLIPOAMIDE ACETYLTRANSFERASE COMPONENT OF PYRUVATE DEHYDROGENASE  
 DE COMPLEX (EC 2.3.1.12) (E2).  
 GN PDHC.  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Staphylococcus.  
 OX NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=V8 / ATCC 27733;  
 RX MEDLINE=92096451; PubMed=1756171;  
 RA Hemila H.;  
 RT "Lipoamide dehydrogenase of Staphylococcus aureus: nucleotide  
 sequence and sequence analysis.";  
 RL Biochim. Biophys. Acta 1129:119-123(1991).  
 CC -1- FUNCTION: THE PYRUVATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL  
 CONVERSION OF PYRUVATE TO ACETYL-COA & CO(2). IT CONTAINS MULTIPLE  
 COPIES OF THREE ENZYMOLOGICAL COMPONENTS: PYRUVATE DEHYDROGENASE (E1),  
 DIHYDROLIPOAMIDE ACETYLTRANSFERASE (E2) & LIPOAMIDE DEHYDROGENASE  
 (E3).  
 CC -1- CATALYTIC ACTIVITY: ACETYL-COA + DIHYDROLIPOAMIDE = COA +  
 S-ACETYLDIHYDROLIPOAMIDE.  
 CC -1- COFACTOR: THE E2 COMPONENT CONTAINS ONE COVALENTLY-BOUND LIPOYL  
 COFACTOR (BY SIMILARITY).  
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL  
 SYMMETRY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 LIPOYL-BINDING DOMAIN  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: X58434; CAA41339.1;  
 DR HSSP: P11961; 2PDE.  
 DR InterPro: IPR000089;  
 DR InterPro: IPR001078;  
 DR InterPro: IPR003016;  
 DR Pfam: PF00198; 2-oxoacid\_dh; 1.  
 DR Pfam: PF00364; biotin\_lioyl; 1.  
 DR PROSITE: PS00189; LIPOYL; 1.  
 KW Glycolysis; Transferrase; Acyltransferase; Lipoyl.  
 FT BINDING 43 43 LIPOYL (BY SIMILARITY).  
 FT ACT\_SITE 401 401 POTENTIAL.  
 SQ SEQUENCE 430 AA; 46439 MW; AFE00748A65ACDA9 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 430;  
 Best Local Similarity 100.0%; Pred. No. 95;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 VDYLNG 1061  
 DB 158 VDYLNG 164

RESULT 70  
 ACRO\_RABIT STANDARD; PRT; 431 AA.  
 ID ACRO\_RABIT  
 AC P48038;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE ACROSIN PRECURSOR (EC 3.4.21.10).  
 GN ACR.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Testis;  
 RX MEDLINE=94368861; PubMed=8086468;  
 RA Richardson R.T.; O'Rand M.G.;  
 RT "Cloning and sequencing of cDNAs for rabbit preproacrosin and a novel  
 preproacrosin-related cDNA.";  
 RL Biochim. Biophys. Acta 1219:215-218(1994).  
 CC -1- FUNCTION: ACROSIN IS THE MAJOR PROTEASE OF MAMMALIAN SPERMATOZOA.  
 IT IS A SERINE PROTEASE OF TRYPSIN-LIKE CLEAVAGE SPECIFICITY. IT  
 IS SYNTHESIZED IN A ZMOGEN FORM, PROACROSIN AND STORED IN THE  
 ACROSOME.  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF ARG- AND LYS-BONDS; PREFERENTIAL  
 CLEAVAGE ARG-XAA >> LYS-LYS >> LYS-XAA.  
 CC -1- SUBUNIT: HEAVY CHAIN (CATALYTIC) AND A LIGHT CHAIN LINKED BY TWO  
 DISULFIDE BONDS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 TRYPSIN FAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: U05204; AAA61630.1;  
 DR MEROPS: S01.223;  
 DR InterPro: IPR001254;  
 DR InterPro: IPR001314;  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS; PS00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Sperm; Signal.  
 FT SIGNAL 1 16  
 FT CHAIN 17 431 ACROSIN.  
 FT CHAIN 17 39 ACROSIN LIGHT CHAIN (BY SIMILARITY).  
 FT CHAIN 40 ? ACROSIN HEAVY CHAIN (BY SIMILARITY).  
 FT PROPEP 2 431 PRO-RICH.  
 FT DISULFID 22 152 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 26 160 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 71 87 BY SIMILARITY.  
 FT DISULFID 175 244 BY SIMILARITY.  
 FT DISULFID 207 223 BY SIMILARITY.  
 FT DISULFID 234 264 BY SIMILARITY.  
 FT CARBOHYD 19 19 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT ACT\_SITE 86 86 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 140 140 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT SEQUENCE 431 AA; 46422 MW; IC015A4E0BC0C668 CRC64;

```

OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE UPF0053 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 295208; CAB08473.1;
DR Tuberculin; RV2366C;
DR InterPro; IPR000644;
DR InterPro; IPR002550;
DR Pfam; PF00571; CBS; 2;
DR Pfam; PF01595; DUF21; 1;
DR KJ000000;
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 7 27
FT TRANSMEM 89 109
FT SEQUENCE 435 AA; 47213 MW; C12E988EC190DD4A CRC64;
SQ
Query Match 0.5%; Score 7; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 889 IESVFEL 895
Db 189 IESVFEL 195
PS31_ARATH
AC Q9LNU4;
ID PS31_ARATH STANDARD; PRT; 488 AA.
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROBABLE 26S PROTEASOME REGULATORY SUBUNIT S3.
GN T20H2.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Sakano H., Vaysberg M., Lee J.M., Lenz C., Liu S., Pham P.,
RA Toriumi M., Yu G., Chin C., Chou J., Choi E., Chung M., Gonzalez A.,
RA Hong B., Liu A., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTS AS A REGULATORY SUBUNIT OF THE 26 PROTEASOME WHICH

```

IS INVOLVED IN THE ATP-DEPENDENT DEGRADATION OF UBIQUITINATED  
PROTEINS (BY SIMILARITY).  
CC -1- SUBUNIT: THE 26S PROTEASOME IS COMPOSED OF A CORE PROTEASE, KNOWN  
AS THE 20S PROTEASOME, CAPPED AT ONE OR BOTH ENDS BY THE 19S  
REGULATORY COMPLEX (RC). THE RC IS COMPOSED OF AT LEAST 18  
DIFFERENT SUBUNITS IN TWO SUBCOMPLEXES, THE BASE AND THE LID,  
WHICH FORM THE PORTIONS PROXIMAL AND DISTAL TO THE 20S PROTEOLYTIC  
CORE, RESPECTIVELY (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE PROTEASOME SUBUNIT S3 FAMILY.  
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; ALTERNATIVE INTRON/EXON  
JUNCTIONS HAVE BEEN USED TO MAXIMIZE SIMILARITIES WITH HOMOLOGS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AC022472; AAF79894.1; ALT\_SEQ.  
KW Hypothetical protein; Proteasome.  
SQ SEQUENCE 488 AA; 55583 MW; 66F0F3F8844300E4 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 488;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 336 NNTPSQS 342  
| | | | |  
DB 10 NNTPSQS 16

RESULT 74  
GAPN\_MAIZE STANDARD; PRT; 498 AA.  
AC Q43272;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE NADP-DEPENDENT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.9)  
DE (NON-PHOSPHORYLATING GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE)  
DE (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE [NADP+]) (TRIOSEPHOSPHATE  
DE DEHYDROGENASE).  
GN GPNI.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;  
OC Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. KW5330; TISSUE=Shoot;  
RX MEDLINE=94180387; PubMed=7545914;  
RA Habenicht A., Hellman U., Cerff R.;  
RT "Non-phosphorylating GAPDH of higher plants is a member of the  
aldehyde dehydrogenase superfamily with no sequence homology to  
phosphorylating GAPDH.";  
RL J. Mol. Biol. 237:165-171(1994).  
CC -1- FUNCTION: IMPORTANT AS A MEANS OF GENERATING NADPH FOR  
BIOSYNTHETIC REACTIONS.  
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + NADP(+) + H(2)O  
= 3-PHOSPHO-D-GLYCERATE + NADPH.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X75326; CAA53075.1; -  
DR MaizetDB; 78926; -  
DR InterPro; IPR002086; -  
DR Pfam; PF001171; aldedh; 1.  
DR PROSITE; PS00070; ALDEHYDE\_DEHYDR\_CYS; 1.  
DR PROSITE; PS00687; ALDEHYDE\_DEHYDR\_GLU; 1.  
KW Oxidoreductase; NADP NAD (ADP PART) (BY SIMILARITY).  
FT NP\_BIND 247 251  
FT ACT\_SITE 247 266 BY SIMILARITY.  
FT ACT\_SITE 300 300 BY SIMILARITY.  
SQ SEQUENCE 498 AA; 53146 MW; 7AF1C0DACAB4EE39 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 498;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1158 NALVLKP 1164  
| | | | |  
DB 189 NALVLKP 195

RESULT 75  
INOL\_CITPA STANDARD; PRT; 507 AA.  
ID INOL\_CITPA  
AC P42802;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE MYO-INOSITOL-1-PHOSPHATE SYNTHASE (EC 5.5.1.4) (IPS).  
OS Citrus paradisi (Grapefruit).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Sapindales; Rutaceae; Citrus.  
OX NCBI\_TaxID=37656;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leaf;  
RX MEDLINE=95148748; PubMed=7846170;  
RA Abu-Abied M., Holland D.;  
RT "The gene c-inol from Citrus paradisi is highly homologous to turl  
RT and inol from yeast and Spirodela encoding for myo-inositol phosphate  
RT synthase.";  
RL Plant Physiol. 106:1689-1689(1994).  
CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE -> 1L-MYO-INOSITOL  
1-PHOSPHATE.  
CC -1- COFACTOR: NAD.  
CC -1- PATHWAY: INOSITOL BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE MYO-INOSITOL-1-PHOSPHATE SYNTHASE  
FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: Z32632; CAA83565.1; -  
DR InterPro; IPR002587; -  
DR Pfam; PF01658; Inos-1-p\_synth; 1.  
KW Phospholipid biosynthesis; Inositol biosynthesis; Isomerase; NAD.  
SQ SEQUENCE 507 AA; 56334 MW; 45D78928991BFDF8 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 507;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
QY 1168 VSYNHLG 1174
Db 333 VSYNHLG 339

RESULT 76
INOL_SPIPO STANDARD; PRT; 510 AA.
AC P42803;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE MYO-INOSITOL-1-PHOSPHATE SYNTHASE (EC 5.5.1.4) (IPS).
GN TUR1.
OS Spirodela polyrrhiza.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Araceae; Spirodela.
OX NCBI_TaxID=29656;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94035182; PubMed=8220483;
RA Smart C.C., Fleming A.J.;
RT "A plant gene with homology to D-myo-inositol-3-phosphate synthase is
RT rapidly and spatially up-regulated during an abscisic-acid-induced
RT morphogenic response in Spirodela polyrrhiza.";
RL Plant J. 4:279-293(1993).
CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = 1L-MYO-INOSITOL
CC 1-PHOSPHATE.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: INOSITOL BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- INDUCTION: BY ABSICISIC ACID (ABA).
CC -1- SIMILARITY: BELONGS TO THE MYO-INOSITOL-1-PHOSPHATE SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z11693; CAA77751.1; -
CC InterPro: IPR002587; -
CC Pfam; PF01658; Inos-1-p-synth; 1.
CC Phospholipid biosynthesis; Inositol biosynthesis; Isomerase; NAD.
CC SEQUENCE 510 AA; 56385 MW; 2D56D3666FC5E03C CRC64;

Query Match 0.5%; Score 7; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174
Db 333 VSYNHLG 339

RESULT 77
INOL_ARATH STANDARD; PRT; 511 AA.
AC P42801;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MYO-INOSITOL-1-PHOSPHATE SYNTHASE (EC 5.5.1.4) (IPS).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=94336770; PubMed=8058832;
RA Johnson M.A.;
RT "The Arabidopsis thaliana myo-inositol 1-phosphate synthase (EC
RT 5.5.1.4).";
RL Plant Physiol. 105:1023-1024(1994).
RN [2]
RP REVISIONS.
RA Johnson M.A.;
RL Submitted (XXX-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = 1L-MYO-INOSITOL
CC 1-PHOSPHATE.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: INOSITOL BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE MYO-INOSITOL-1-PHOSPHATE SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U04876; AAA85390.1; -
CC HSPP; P10246; 1TRF.
CC InterPro: IPR002587; -
CC Pfam; PF01658; Inos-1-P-synth; 1.
CC Phospholipid biosynthesis; Inositol biosynthesis; Isomerase; NAD.
CC SEQUENCE 511 AA; 56444 MW; 0D4837D372D3B579 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 511;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174
Db 334 VSYNHLG 340

RESULT 78
SPI_RARFA STANDARD; PRT; 525 AA.
AC Q05308;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SERINE PROTEASE I PRECURSOR (EC 3.4.21.-) (RPI).
OS Rarobacter faecitabidus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococineae; Rarobacteraceae; Rarobacter.
OX NCBI_TaxID=13243;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 212-222; 224-238 AND 240-244.
RC STRAIN=YLM-50;
RX MEDLINE=93094226; PubMed=1339445;
RA Shimoi H., Iimura Y., Obata T., Tadenuma M.;
RT "Molecular structure of Rarobacter faecitabidus protease I. A yeast-
RT lytic serine protease having mannose-binding activity.";
RL J. Biol. Chem. 267:25189-25195(1992).
RN [2]
RP SEQUENCE OF 212-247.
RX MEDLINE=92138668; PubMed=1778983;
RA Shimoi H., Tadenuma M.;
RT "Characterization of Rarobacter faecitabidus protease I, a
RT yeast-lytic serine protease having mannose-binding activity.";
RL J. Biochem. 110:608-613(1991).
CC -1- FUNCTION: MAJOR SERINE PROTEASE EXHIBITING LYTIC ACTIVITY TOWARD
CC LIVING YEAST CELLS. SIMILAR TO ELASTASE IN ITS SUBSTRATE
```



SPECIFICITY AND HAS A LECTIN-LIKE AFFINITY FOR MANNOSE.  
MANNOSE-BINDING DOMAIN.  
-|- DOMAIN: COMPOSED OF N-TERMINAL PROTEASE DOMAIN AND A C-TERMINAL  
MANNOSE-BINDING DOMAIN.  
-|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2A; ALSO KNOWN AS THE  
ALPHA-LYTIC PROTEASE FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: D10753; BAA01585.1; -  
PIR: A45053; A45053.  
HSP: P00778; ITAL.  
MEROPS: S01.276; -  
InterPro: IPR000772; -  
InterPro: IPR001254; -  
InterPro: IPR001316; -  
Pfam: PF00652; Ricin\_B\_lectin; 1.  
PRINTS: PR00089; trypsin; 1.  
PROSITE: PS00134; TRYPSIN\_HIS; 1.  
PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Hydrolyase; Serine protease; Mannose-binding; Signal; Zymogen.  
FT SIGNAL 1 32 POTENTIAL.  
FT PROPEP 33 211 POTENTIAL.  
FT CHAIN 212 525 SERINE PROTEASE I.  
FT ACT\_SITE 238 525 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 270 270 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 352 352 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 223 239 BY SIMILARITY.  
FT DISULFID 310 320 BY SIMILARITY.  
FT DISULFID 346 376 BY SIMILARITY.  
FT DISULFID 412 431 BY SIMILARITY.  
FT DISULFID 453 472 BY SIMILARITY.  
FT DISULFID 496 514 BY SIMILARITY.  
FT DOMAIN 412 473 MANNOSE-BINDING (BY SIMILARITY).  
FT DOMAIN 401 525 ESSENTIAL FOR THE LYTIC ACTIVITY, BUT NOT  
FOR PROTEASE FUNCTION.  
SQ SEQUENCE 525 AA; 55654 MW; DA2BCF7D330EBB61 CRC64;  
  
Query Match 0.5%; Score 7; DB 1; Length 525;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 17 LALVGAL 23  
Db 13 LALVGAL 19  
[1]  
  
RESULT 79  
YMOO\_YEAST  
ID YMOO\_YEAST STANDARD; PRT; 532 AA.  
AC Q0458;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL ALDEHYDE-DEHYDROGENASE LIKE PROTEIN IN ILV2-ADE17  
DE INTERGENIC REGION.  
GN YMR110C OR YMR718.09C  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;

Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;  
Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
-|- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: Z45702; CAA89746.1; -  
HSP: P11883; IAD3.  
SGD: S0004716; YMR110C.  
InterPro: IPR002086; -  
Pfam: PF00171; aldedh; 1.  
PROSITE: PS00070; ALDEHYDE DEHYDR\_CYS; FALSE\_NEG.  
PROSITE: PS00687; ALDEHYDE DEHYDR\_GLU; 1.  
KW Hypothetical protein; Oxidoreductase.  
FT ACT\_SITE 236 236 BY SIMILARITY.  
FT ACT\_SITE 273 273 BY SIMILARITY.  
SQ SEQUENCE 532 AA; 59978 MW; 6CAF4BFCF963AF88 CRC64;  
  
Query Match 0.5%; Score 7; DB 1; Length 532;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 371 PPNSAQK 377  
Db 480 PPNSAQK 486  
[1]  
  
RESULT 80  
HEXB\_MOUSE  
ID HEXB\_MOUSE STANDARD; PRT; 536 AA.  
AC P20060;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE BETA-HEXOSAMINIDASE BETA CHAIN PRECURSOR (PC 3.2.1.52) (N-ACETYL-BETA-  
GLUCOSAMINIDASE) (BETA-N-ACETYLHEXOSAMINIDASE) (HEXOSAMINIDASE A).  
GN HEXB.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89005625; PubMed=2971567;  
RA Bapat B., Ethier M., Neote K., Mahuran D., Gravel R.A.;  
RT "Cloning and sequence analysis of a cDNA encoding the beta-subunit of  
mouse beta-hexosaminidase";  
RL FEBS Lett. 237:191-195(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B6/CBA; TISSUE=Liver;  
RX MEDLINE=95048337; PubMed=7959736;  
RA Yamanaka S., Johnson O.N., Norflus F., Boles D.J., Proia R.L.;  
RT "Structure and expression of the mouse beta-hexosaminidase genes,  
Hexa and Hexb.";  
RL Genomics 21:588-596(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RX MEDLINE=95002207; PubMed=7918686;  
RA Triggs-Raine B.L., Benoit G., Salo T.J., Trasler J.M., Gravel R.A.;  
RT "Characterization of the murine beta-hexosaminidase (HEXB) gene";  
RL Biochim. Biophys. Acta 1227:79-86(1994).  
CC -|- FUNCTION: BETA-HEXOSAMINIDASE A IS RESPONSIBLE FOR THE DEGRADATION  
OF GM2 GANGLIOSIDES, AND A VARIETY OF OTHER MOLECULES CONTAINING  
TERMINAL N-ACETYL HEXOSAMINES, IN THE BRAIN AND OTHER TISSUES.  
CC

CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING N-ACETYL-  
 CC D-HEXOSAMINE RESIDUES IN N-ACETYL-BETA-D-HEXOSAMINIDES  
 CC -1- SUBUNIT: THERE ARE 3 FORMS OF BETA-HEXOSAMINIDASE: HEXOSAMINIDASE  
 CC A IS A TRIMER COMPOSED OF ONE ALPHA CHAIN, ONE BETA-A CHAIN & ONE  
 CC BETA-B CHAIN. HEXOSAMINIDASE B IS A TETRAMER OF TWO BETA-A AND TWO  
 CC BETA-B CHAINS. HEXOSAMINIDASE S IS AN HOMODIMER OF TWO ALPHA  
 CC CHAINS. THE TWO BETA CHAINS ARE DERIVED FROM THE CLEAVAGE OF A  
 CC PRECURSOR CHAIN (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 20 OF GLYCOSYL HYDROLASES. STRONG,  
 CC TO ALPHA CHAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Y00964; CAAB8781.1; -  
 CC EMBL; U07633; AAA18776.1; -  
 CC EMBL; U07049; AAA74738.1; -  
 CC EMBL; U07036; AAA74738.1; JOINED.  
 CC EMBL; U07037; AAA74738.1; JOINED.  
 CC EMBL; U07038; AAA74738.1; JOINED.  
 CC EMBL; U07039; AAA74738.1; JOINED.  
 CC EMBL; U07040; AAA74738.1; JOINED.  
 CC EMBL; U07041; AAA74738.1; JOINED.  
 CC EMBL; U07042; AAA74738.1; JOINED.  
 CC EMBL; U07043; AAA74738.1; JOINED.  
 CC EMBL; U07044; AAA74738.1; JOINED.  
 CC EMBL; U07045; AAA74738.1; JOINED.  
 CC EMBL; U07046; AAA74738.1; JOINED.  
 CC EMBL; U07047; AAA74738.1; JOINED.  
 CC EMBL; U07048; AAA74738.1; JOINED.  
 CC EMBL; U07049; AAA74738.1; JOINED.  
 CC EMBL; U07042; AAB60667.1; -  
 CC EMBL; U07722; AAB60667.1; JOINED.  
 CC EMBL; U07723; AAB60667.1; JOINED.  
 CC EMBL; U07724; AAB60667.1; JOINED.  
 CC EMBL; U07725; AAB60667.1; JOINED.  
 CC EMBL; U07726; AAB60667.1; JOINED.  
 CC EMBL; U07727; AAB60667.1; JOINED.  
 CC EMBL; U07728; AAB60667.1; JOINED.  
 CC EMBL; U07737; AAB60667.1; JOINED.  
 CC EMBL; U07738; AAB60667.1; JOINED.  
 CC EMBL; U07739; AAB60667.1; JOINED.  
 CC EMBL; U07740; AAB60667.1; JOINED.  
 CC EMBL; U07741; AAB60667.1; JOINED.  
 CC PIR; S01328; S01328.  
 CC HSP; P07686; IQBD.  
 CC MGD; MGI:96074; Hexb.  
 CC InterPro: IPR001540; -  
 CC Pfam: PF00728; Glyco\_hydro\_20; 1.  
 CC PRINTS; PR00738; GLYDRLASE20.  
 CC KW Hydrolyase; Glycosidase; Lysosome; Signal; zymogen; Glycoprotein.  
 CC FT SIGNAL 1 24  
 CC FT PROPEP 25 ?  
 CC FT CHAIN ? 536 BETA-HEXOSAMINIDASE BETA CHAIN.  
 CC FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT ACT\_SITE 334 334 CATALYTIC ACID (BINDS TO THE GLYCOSIDIC  
 CC LINKAGE) (BY SIMILARITY).  
 CC SEQUENCE 536 AA; 61115 MW; 579BBEE9CB508BC CRC64;

Query Match 0.5%; Score 7; DB 1; Length 536;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LVSLALV 20  
 |||||

Db 20 LVSLALV 26

RESULT 81  
 ID ESTJ\_HELVI STANDARD; PRT; 564 AA.  
 AC P12992;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE JUVENILE HORMONE ESTERASE PRECURSOR (EC 3.1.1.59) (JH ESTERASE).  
 OS Heliothis virescens (Noctuid moth) (Owlet moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Noctuoidea; Noctuidae; Heliothinae; Heliothis.  
 OX NCBI\_TaxID=7102;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-54.  
 RX MEDLINE=89308671; PubMed=2745451;  
 RA Hanzlik T.N., Yehia A.I.A.-A., Harshman L.G., Hammock B.D.;  
 RT "Isolation and sequencing of cDNA clones coding for juvenile hormone  
 RT esterase from *Heliothis virescens*. Evidence for a catalytic mechanism  
 RT for the serine carboxylesterases different from that of the serine  
 RT proteases.";  
 RL J. Biol. Chem. 264:12419-12425(1989).  
 RN [2]  
 RP REVISIONS.  
 RA Hanzlik T.N.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: JH ESTERASE PLAYS A CRUCIAL ROLE IN THE DECREASE OF  
 CC JH ACTIVITY IN LEPIDOPTERAN INSECTS BY HYDROLYZING THE METHYL  
 CC ESTER OF JH. IT IS ALSO INVOLVED IN THE TRANSPORT OF JH.  
 CC -1- CATALYTIC ACTIVITY: METHYL (2E,6E)-(10R,11S)-10,11-EPOXY-3,7,11-  
 CC TRIMETHYLTRIDECA-2,6-DIENOATE + H(2)O = (2E,6E)-(10R,11S)-10,11-  
 CC EPOXY-3,7,11-TRIMETHYLTRIDECA-2,6-DIENOATE + METHANOL.  
 CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; J04955; AAB88629.1; -  
 CC PIR; A34325; A34325.  
 CC HSP; P21836; IMAH.  
 CC InterPro: IPR002018; -  
 CC Pfam; PF00135; Coesterase; 1.  
 CC PROSITE; PS00122; CARBOXYLESTERASE\_B\_1; 1.  
 CC PROSITE; PS00941; CARBOXYLESTERASE\_B\_2; FALSE NEG.  
 CC KW Hydrolyase; Serine esterase; Glycoprotein; Signal.  
 CC FT SIGNAL 1 19  
 CC FT CHAIN 20 564 JUVENILE HORMONE ESTERASE.  
 CC FT ACT\_SITE 220 220 BY SIMILARITY.  
 CC FT ACT\_SITE 351 351 BY SIMILARITY.  
 CC FT ACT\_SITE 465 465 BY SIMILARITY.  
 CC FT DISULFID 89 109 BY SIMILARITY.  
 CC FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT VARIANT 29 29 V -> L.  
 CC FT VARIANT 52 52 F -> P.  
 CC SEQUENCE 564 AA; 62614 MW; D140E5DD91914E8D CRC64;

Query Match 0.5%; Score 7; DB 1; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 ISIDNFV 201

```
Db 405 ISIDNEV 411
|||||
RESULT 82
ID IF2_THETH STANDARD; PRT; 571 AA.
AC P48515;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TRANSLATION INITIATION FACTOR IF-2.
GN INF2.
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=HB8 / ATCC 27634;
RA Vornlocher H.P., Scheible W.R., Faulhammer H.G., Sprinzl M.;
RT "Identification and purification of translation initiation factor 2 (IF2) from Thermus thermophilus.";
RL Eur. J. Biochem. 243:66-71(1997).
CC -!- FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION OF PROTEIN SYNTHESIS IN VITRO. PROTECTS FORMYLMETHIONYL-TRNA FROM SPONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S RIBOSOMAL SUBUNIT. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z48001; CAA88038.1; -
DR HSP; P07157; 1AIP.
DR InterPro; IPR000178; -
DR InterPro; IPR000795; -
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF02131; IF2; 1.
DR PROSITE; PS01176; IF2; 1.
KW Initiation factor; Protein biosynthesis; GTP-binding.
FT DOMAIN 72 222 G-DOMAIN
FT NP_BIND 80 87 GTP (BY SIMILARITY).
FT NP_BIND 126 130 GTP (BY SIMILARITY).
FT NP_BIND 180 183 GTP (BY SIMILARITY).
SQ SEQUENCE 571 AA; 63130 MW; 7004210C19C90F01 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 571;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 QAGIIAN 305
Db 259 QAGIIAN 265
|||||
RESULT 83
SKOL_YEAST
ID SKOL_YEAST STANDARD; PRT; 647 AA.
AC Q02100;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE CRE-BINDING BZIP PROTEIN SKOL.
GN SKOL OR ACR1 OR INL167C OR NI702.

Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=W303-1A;
RA Nehlin J.O., Carlberg M., Ronne H.;
RT "Yeast SKOL gene encodes a BZIP protein that binds to the CRE motif and acts as a repressor of transcription.";
RL Nucleic Acids Res. 20:5271-5278(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93078739; PubMed=1448073;
RA Vincent A.C., Struhl K.;
RT "ACR1, a yeast ATF/CREB repressor.";
RL Mol. Cell. Biol. 12:5394-5405(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / FY1679;
RA Nasr F., Becam A.-M., Herbert C.J.;
RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals 24 complete open reading frames; 18 correspond to new genes, one of which encodes a protein similar to the human myotonic dystrophy kinase.";
RL Yeast 12:169-175(1996).
CC -!- FUNCTION: BINDS TO THE CRE MOTIF 5'-TGACGTCA-3' AND ACTS AS A REPRESSOR OF TRANSCRIPTION OF THE SUC2 GENE AND MOST PROBABLY OTHER GENES.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: TO OTHER BZIP PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67875; CAA48074.1; -
DR EMBL; S49588; AAB24288.1; -
DR EMBL; X92517; CAA63272.1; -
DR EMBL; Z71443; CAA96054.1; -
DR PIR; S26386; S26386.
DR PIR; A45028; A45028.
DR HSP; P05412; 1FOS.
DR TRANSFAC; T01306; -
DR SGD; S0005111; SKOL.
DR InterPro; IPR001871; -
DR Pfam; PF00170; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; FALSE NEG.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DNA_BIND 431 456 BASIC MOTIF.
FT DOMAIN 457 478 LEUCINE-ZIPPER.
SQ SEQUENCE 647 AA; 70192 MW; 3E0B8C72A6CE14AB CRC64;

Query Match 0.5%; Score 7; DB 1; Length 647;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1084 SLNSGAN 1090
Db 588 SLNSGAN 594
|||||
RESULT 84
Y4ID_RHISN
ID Y4ID_RHISN STANDARD; PRT; 662 AA.
AC P55487;
```

DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PROBABLE MONOOXYGENASE Y4ID (EC 1.14.13.-).  
 GN Y4ID.  
 OS Rhizobium sp. (strain NGR234).  
 OG Plasmid sym pNGR234a.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97305956; PubMed=9163424;  
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,  
 RA Perret X.;  
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";  
 RL Nature 387:394-401(1997).  
 CC -1- SIMILARITY: TO ACINETOBACTER SP. CYCLOHEXANONE MONOOXYGENASE,  
 CC -1- COFACTOR: FAD (BY SIMILARITY).  
 CC M.TUBERCULOSIS RV0892, A.NIDULANS STCW AND TO MAMMALIAN  
 CC DIMETHYLANILINE MONOOXYGENASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: AE000078; AB91699.1; -  
 CC InterPro: IPR000960; -  
 CC Pfam: PF00743; FMO-like; 1.  
 CC Hypothetical protein; Oxidoreductase; Monooxygenase; NADP;  
 KW Flavoprotein; FAD; plasmid.  
 FT NP BIND 303 335 NADP (BY SIMILARITY).  
 SQ SEQUENCE 562 AA; 75109 MW; 524689EA62AA76CB CRC64;

Query Match 0.5%; Score 7; DB 1; Length 662;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LSLGRNF 151  
 |||||  
 Db 278 LSLGRNF 284

RESULT 85  
 CICK\_RABIT STANDARD; PRT; 678 AA.  
 ID CICK\_RABIT STANDARD; PRT; 678 AA.  
 AC P51804;  
 DT 01-OCT-1996 (Rel. 34, Created).  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CHLORIDE CHANNEL PROTEIN CLC-K2.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96163242; PubMed=8587242;  
 RA Zimniak L., Winters C.J., Reeves W.B., Andreoli T.E.;  
 RT "Cl- channels in basolateral renal medullary vesicles. X. Cloning of  
 RT a Cl- channel from rabbit outer medulla.";  
 RL Kidney Int. 48:1828-1836(1995).  
 CC -1- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE  
 CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;  
 CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND  
 CC TRANSEPITHelial TRANSPORT. MAY BE IMPORTANT IN URINARY  
 CC CONCENTRATING MECHANISMS.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE KIDNEY.  
 CC -1- SIMILARITY: TO OTHER CHLORIDE CHANNELS.  
 CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: U36790; AAC48493.1; -  
 CC InterPro: IPR000644; -  
 CC InterPro: IPR001807; -  
 CC InterPro: IPR002250; -  
 CC Pfam: PF00571; CBS; 2.  
 CC Pfam: PF00654; Voltage\_CLC; 1.  
 CC PRINTS: PRO0762; CLCHANNEL.  
 CC PRINTS: PRO1119; CLCHANNELKDY.  
 CC Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;  
 KW CBS domain.  
 FT DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 50 69 1 (POTENTIAL).  
 FT TRANSMEM 91 114 2 (POTENTIAL).  
 FT TRANSMEM 138 159 3 (POTENTIAL).  
 FT TRANSMEM 169 188 4 (POTENTIAL).  
 FT TRANSMEM 200 224 5 (POTENTIAL).  
 FT TRANSMEM 239 257 6 (POTENTIAL).  
 FT TRANSMEM 282 302 7 (POTENTIAL).  
 FT TRANSMEM 325 348 8 (POTENTIAL).  
 FT TRANSMEM 398 417 9 (POTENTIAL).  
 FT TRANSMEM 420 438 10 (POTENTIAL).  
 FT TRANSMEM 466 487 11 (POTENTIAL).  
 FT TRANSMEM 495 514 12 (POTENTIAL).  
 FT DOMAIN 515 545 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 546 564 13 (POTENTIAL).  
 FT DOMAIN 565 678 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 549 604 CBS 1.  
 FT DOMAIN 620 678 CBS 2.  
 FT CARBOHYD 364 364 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 678 AA; 74468 MW; 62816AB2877125F2 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 678;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 467 GGYALAG 473  
 |||||  
 Db 464 GGYALAG 470

RESULT 86  
 CICK\_HUMAN STANDARD; PRT; 687 AA.  
 ID CICK\_HUMAN STANDARD; PRT; 687 AA.  
 AC P51800;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CHLORIDE CHANNEL PROTEIN CLC-K1.  
 GN CLCNKA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94316614; PubMed=8041726;  
 RA Kieferle S., Fong P., Bens M., Vandewalle A., Jentsch T.;  
 RT "Two highly homologous members of the ClC chloride channel family in

both rat and human kidney.";  
 Proc. Natl. Acad. Sci. U.S.A. 91:6943-6947(1994).  
 -I- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE  
 SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;  
 MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND  
 TRANSEPITHELIAL TRANSPORT. MAY BE IMPORTANT IN URINARY  
 CONCENTRATING MECHANISMS.  
 -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 -I- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE KIDNEY.  
 -I- SIMILARITY: TO OTHER CHLORIDE CHANNELS.  
 -I- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: Z30643; CAA83120.1; -  
 MIM: 602024; -  
 InterPro: IPR000644; -  
 InterPro: IPR001807; -  
 InterPro: IPR002250; -  
 Pfam: PF00571; CBS; 2;  
 Pfam: PF00654; voltage\_CLC; 1;  
 PRINTS; PR00762; CLCHANNEL.  
 PRINTS; PR01119; CLCHANNELKDY.  
 Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;  
 CBS domain.  
 DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).  
 TRANSMEM 50 69 1 (POTENTIAL).  
 TRANSMEM 91 114 2 (POTENTIAL).  
 TRANSMEM 138 159 3 (POTENTIAL).  
 TRANSMEM 169 188 4 (POTENTIAL).  
 TRANSMEM 200 224 5 (POTENTIAL).  
 TRANSMEM 239 257 6 (POTENTIAL).  
 TRANSMEM 282 302 7 (POTENTIAL).  
 TRANSMEM 325 348 8 (POTENTIAL).  
 TRANSMEM 398 417 9 (POTENTIAL).  
 TRANSMEM 420 438 10 (POTENTIAL).  
 TRANSMEM 466 487 11 (POTENTIAL).  
 TRANSMEM 495 514 12 (POTENTIAL).  
 DOMAIN 515 645 CYTOPLASMIC (POTENTIAL).  
 TRANSMEM 646 664 13 (POTENTIAL).  
 DOMAIN 665 687 EXTRACELLULAR (POTENTIAL).  
 DOMAIN 620 678 CBS 1.  
 DOMAIN 679 679 CBS 2.  
 CARBOHYD 679 679 N-LINKED (GLCNAC...) (POTENTIAL).  
 SEQUENCE 687 AA; 75284 MW; 597C6928470A4460 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 687;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 467 GGYALAG 473  
 Db 464 GGYALAG 470

RESULT 87  
 CLCK\_RABIT STANDARD; PRT; 687 AA.  
 AC P51803;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CHLORIDE CHANNEL PROTEIN CLC-K1.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

NCBI\_TaxID=9986;  
 [1]  
 SEQUENCE FROM N.A.  
 TISSUE=Kidney;  
 RX MEDLINE=96163242; PubMed=8587242;  
 RA Zimlak L., Winters C.J., Reeves W.B., Andreoli T.E.;  
 RT "Cl- channels in basolateral renal medullary vesicles. X. Cloning of  
 a Cl- channel from rabbit outer medulla.";  
 RL Kidney Int. 48:1828-1836(1995).  
 CC -I- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE  
 SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;  
 CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND  
 CC TRANSEPITHELIAL TRANSPORT. MAY BE IMPORTANT IN URINARY  
 CC CONCENTRATING MECHANISMS.  
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -I- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE KIDNEY.  
 CC -I- SIMILARITY: TO OTHER CHLORIDE CHANNELS.  
 CC -I- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: U36789; AAC48492.1; -  
 InterPro: IPR000644; -  
 InterPro: IPR001807; -  
 InterPro: IPR002250; -  
 Pfam: PF00571; CBS; 2;  
 Pfam: PF00654; voltage\_CLC; 1;  
 PRINTS; PR00762; CLCHANNEL.  
 PRINTS; PR01119; CLCHANNELKDY.  
 Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;  
 CBS domain.  
 DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).  
 TRANSMEM 50 69 1 (POTENTIAL).  
 TRANSMEM 91 114 2 (POTENTIAL).  
 TRANSMEM 138 159 3 (POTENTIAL).  
 TRANSMEM 169 188 4 (POTENTIAL).  
 TRANSMEM 200 224 5 (POTENTIAL).  
 TRANSMEM 239 257 6 (POTENTIAL).  
 TRANSMEM 282 302 7 (POTENTIAL).  
 TRANSMEM 325 348 8 (POTENTIAL).  
 TRANSMEM 398 417 9 (POTENTIAL).  
 TRANSMEM 420 438 10 (POTENTIAL).  
 TRANSMEM 466 487 11 (POTENTIAL).  
 TRANSMEM 495 514 12 (POTENTIAL).  
 DOMAIN 515 645 CYTOPLASMIC (POTENTIAL).  
 TRANSMEM 646 664 13 (POTENTIAL).  
 DOMAIN 665 687 EXTRACELLULAR (POTENTIAL).  
 DOMAIN 549 604 CBS 1.  
 DOMAIN 620 678 CBS 2.  
 CARBOHYD 364 364 N-LINKED (GLCNAC...) (POTENTIAL).  
 SEQUENCE 687 AA; 75201 MW; B4BD87A43078E28E CRC64;

Query Match 0.5%; Score 7; DB 1; Length 687;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 467 GGYALAG 473  
 Db 464 GGYALAG 470

RESULT 88  
 CLCK\_RAT STANDARD; PRT; 687 AA.  
 ID CLCK\_RAT  
 AC Q06393; P97709;  
 DT 01-JUN-1994 (Rel. 29, Created)

15-JUL-1998 (Rel. 36, Last sequence update)  
 15-JUL-1998 (Rel. 36, Last annotation update)  
 CLORIDE CHANNEL PROTEIN CLC-K1.  
 Rattus norvegicus (Rat).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 [1]  
 SEQUENCE FROM N.A.  
 RP TISSUE=Kidney;  
 RC MEDLINE=93179373; PubMed=7680033;  
 RA Uchida S., Sasaki S., Furukawa T., Hiraoka M., Imai T., Hirata Y.,  
 Marumo F.;  
 RA "Molecular cloning of a chloride channel that is regulated by  
 dehydration and expressed predominantly in kidney medulla.";  
 RL J. Biol. Chem. 268:3821-3824(1993).  
 [2]  
 RP REVISIONS.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;  
 RX MEDLINE=94308189; PubMed=8034678;  
 RA Uchida S., Sasaki S., Furukawa T., Hiraoka M., Imai T., Hirata Y.,  
 Marumo F.;  
 RA "Molecular cloning of a chloride channel that is regulated by  
 dehydration and expressed predominantly in kidney medulla.";  
 RL J. Biol. Chem. 269:19192-19192(1994).  
 [3]  
 RP REVISIONS.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;  
 RA Uchida S.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=94316614; PubMed=8041726;  
 RA Kieferle S., Fong P., Bens M., Vandewalle A., Jentsch T.;  
 RA "Two highly homologous members of the CLC chloride channel family in  
 both rat and human kidney.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6943-6947(1994).  
 CC -1- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE  
 SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;  
 MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND  
 TRANSEPITHELIAL TRANSPORT. MAY BE IMPORTANT IN URINARY  
 CONCENTRATING MECHANISMS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE KIDNEY.  
 CC EXPRESSED STRONGLY IN THE CORTICAL THICK ASCENDING LIMB AND THE  
 DISTAL CONVOLUTED TUBULE, WITH MINOR EXPRESSION IN THE S3 SEGMENT  
 OF THE PROXIMAL TUBULE AND THE CORTICAL COLLECTING TUBULE.  
 CC -1- SIMILARITY: TO OTHER CHLORIDE CHANNELS.  
 CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D13927; BAA03026.1; -;  
 DR EMBL; Z34291; CAA84064.1; -;  
 DR PIR; A45483; A45483.  
 DR InterPro; IPR000644; -;  
 DR InterPro; IPR001807; -;  
 DR InterPro; IPR002250; -;  
 DR Pfam; PF00571; CBS; 2;  
 DR Pfam; PF00654; voltage\_CLC; 1;  
 DR PRINTS; PR00762; CLCHANNEL.  
 DR PRINTS; PR01119; CLCHANNELKDY.  
 KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;  
 KW CBS domain. 1 49 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 50 69 1 (POTENTIAL).  
 FT TRANSMEM 50 69 1 (POTENTIAL).

FT TRANSMEM 91 114 2 (POTENTIAL).  
 FT TRANSMEM 138 159 3 (POTENTIAL).  
 FT TRANSMEM 169 188 4 (POTENTIAL).  
 FT TRANSMEM 200 224 5 (POTENTIAL).  
 FT TRANSMEM 239 257 6 (POTENTIAL).  
 FT TRANSMEM 282 302 7 (POTENTIAL).  
 FT TRANSMEM 325 348 8 (POTENTIAL).  
 FT TRANSMEM 398 417 9 (POTENTIAL).  
 FT TRANSMEM 420 438 10 (POTENTIAL).  
 FT TRANSMEM 466 487 11 (POTENTIAL).  
 FT TRANSMEM 495 514 12 (POTENTIAL).  
 FT DOMAIN 515 645 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 646 664 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 665 687 CBS 1.  
 FT DOMAIN 687 704 CBS 2.  
 FT DOMAIN 704 720 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 364 364 I -> V (IN REF. 4).  
 FT CONFLICT 266 266 W -> R (IN REF. 4).  
 FT CONFLICT 534 534 AS -> TP (IN REF. 4).  
 FT CONFLICT 608 609  
 SQ SEQUENCE 687 AA; 75569 MW; 41434F07E3E6E8AD CRC64;

Query Match 0.5%; Score 7; DB 1; Length 687;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 467 GGYALAG 473  
 |||||  
 Db 464 GGYALAG 470

RESULT 89  
 CICL HUMAN STANDARD; PRT; 687 AA.  
 ID CICL\_HUMAN  
 AC P51801;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CHLORIDE CHANNEL PROTEIN CLC-KB (CLC-K2).  
 GN CLCNKB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=94316614; PubMed=8041726;  
 RA Kieferle S., Fong P., Bens M., Vandewalle A., Jentsch T.;  
 RA "Two highly homologous members of the CLC chloride channel family in  
 both rat and human kidney.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6943-6947(1994).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=96130539; PubMed=8544406;  
 RA Takeuchi Y., Uchida S., Marumo F., Sasaki S.;  
 RA "Cloning, tissue distribution, and intrarenal localization of CLC  
 chloride channels in human kidney.";  
 RL Kidney Int. 48:1497-1503(1995).  
 [3]  
 RP VARIANTS BARTTER SYNDROME L-124; T-204; D-349; H-432 AND C-438.  
 RX MEDLINE=97467727; PubMed=9326936;  
 RA Simon D.B., Bindra R.S., Mansfield T.A., Nelson-Williams C.,  
 Mendonca E., Stone R., Schurman S., Nayir A., Alpay H., Bakaloglu A.,  
 Rodriguez-Soriano J., Morales J.M., Sanjad S.A., Taylor C.M.,  
 Pilz D., Brem A., Trachtman H., Griswold W., Richard G.A., John E.,  
 Lifton R.P.;  
 RA "Mutations in the chloride channel gene, CLCNKB, cause Bartter's  
 syndrome type III.";  
 RL Nat. Genet. 17:171-178(1997).  
 CC -1- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE

SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;  
MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND  
TRANSEPITHELIAL TRANSPORT. MAY BE IMPORTANT IN URINARY  
CONCENTRATING MECHANISMS.

-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

-1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE KIDNEY.

-1- DISEASE: DEFECTS IN CLCNKB ARE A CAUSE OF TYPE III BARTTER  
SYNDROME, AN AUTOSOMAL RECESSIVE FORM OF OFTEN SEVERE  
INTRAVASCULAR VOLUME DEPLETION DUE TO RENAL SALT-WASTING  
ASSOCIATED WITH LOW BLOOD PRESSURE, HYPOKALEMIC ALKALOSIS,  
HYPERCALCIURIA, AND NORMAL SERUM MAGNESIUM LEVELS.

-1- SIMILARITY: TO OTHER CHLORIDE CHANNELS.

-1- SIMILARITY: CONTAINS 2 CBS DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; Z30644; CAA83121.1; -

EMBL; S80315; AAB35898.1; -

MIM; 602023; -

MIM; 241200; -

InterPro; IPR000644; -

InterPro; IPR001807; -

InterPro; IPR002250; -

Pfam; PF00571; CBS; 2;

Pfam; PF00654; voltage\_CLC; 1;

PRINTS; PR00762; CLCHANNEL.

PRINTS; PR01119; CLCHANNELKDY.

Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;

CBS domain; Disease mutation.

DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).

1 (POTENTIAL).

TRANSMEM 50 69 2 (POTENTIAL).

TRANSMEM 91 114 3 (POTENTIAL).

TRANSMEM 138 159 4 (POTENTIAL).

TRANSMEM 169 188 5 (POTENTIAL).

TRANSMEM 200 224 6 (POTENTIAL).

TRANSMEM 239 257 7 (POTENTIAL).

TRANSMEM 282 302 8 (POTENTIAL).

TRANSMEM 325 348 9 (POTENTIAL).

TRANSMEM 398 417 10 (POTENTIAL).

TRANSMEM 420 438 11 (POTENTIAL).

TRANSMEM 466 487 12 (POTENTIAL).

TRANSMEM 495 514 13 (POTENTIAL).

DOMAIN 515 645 CYTOPLASMIC (POTENTIAL).

TRANSMEM 646 664 13 (POTENTIAL).

TRANSMEM 665 687 EXTRACELLULAR (POTENTIAL).

DOMAIN 549 604 CBS 1.

DOMAIN 620 678 CBS 2.

DOMAIN 679 699 N-LINKED (GLCNAC. . .) (POTENTIAL).

VARIANT 124 124 P -> L (IN BARTTER SYNDROME TYPE 3).

FT VARIANT 204 204 A -> T (IN BARTTER SYNDROME TYPE 3).

FT VARIANT 349 349 A -> D (IN BARTTER SYNDROME TYPE 3).

FT VARIANT 432 432 Y -> H (IN BARTTER SYNDROME TYPE 3).

FT VARIANT 438 438 R -> C (IN BARTTER SYNDROME TYPE 3).

FT SEQUENCE 687 AA; 75387 MW; 35DDF654B3BD70A4 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 687;

Best Local Similarity 100.0%; Pred. No. 1.4e-02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 467 GGYALAG 473

|||||

Db 464 GGYALAG 470

RESULT 90

PERE\_MOUSE STANDARD; PRT; 716 AA.

AC P49290; Q61798;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE EOSINOPHIL PEROXIDASE PRECURSOR (EC 1.11.1.7) (EPO).

GN EPX OR EPER.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6;

RA Ohmori J., Itoh H., Tomita M., Nawa Y.;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Bone marrow;

RX MEDLINE=96369651; PubMed=8773591;

RA Horton M.A., Larson K.A., Lee J.J., Lee N.A.;

RT "Cloning of the murine eosinophil peroxidase gene (mEPO):

RT characterization of a conserved subgroup of mammalian hematopoietic

RT peroxidases.";

RL J. Leukoc. Biol. 60:285-294(1996).

CC -1- CATALYTIC ACTIVITY: DONOR + H(2)O(2) = OXIDIZED DONOR + 2 H(2)O.

CC -1- COFACTOR: HEME (PROTOPHYRIN IX).

CC -1- SUBUNIT: Tetramer of two light chains and two heavy chains (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES OF EOSINOPHILS.

CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. XPO SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; D78353; BAA1370.1; -

EMBL; L77979; AAB40403.1; ALT\_INIT.

HSP; P05164; IMHL.

MGD; MGI:107569; Epx.

InterPro; IPR002007; -

InterPro; IPR002016; -

Pfam; PF00141; peroxidase; 1.

PRINTS; PR00457; ANPEROXIDASE.

PROSITE; PS00435; PEROXIDASE\_1; 1.

PROSITE; PS00436; PEROXIDASE\_2; FALSE\_NEG.

Oxidoreductase; Peroxidase; Heme; Glycoprotein; Eosinophil; Signal.

SIGNAL 1 18 POTENTIAL.

PROPEP 19 140 POTENTIAL.

CHAIN 141 251 LIGHT CHAIN.

CHAIN 252 716 HEAVY CHAIN.

ACT\_SITE 234 234 DISTAL HISTIDINE (POTENTIAL).

ACT\_SITE 378 378 DISTAL ARGININE (POTENTIAL).

ACT\_SITE 475 475 PROXIMAL HISTIDINE (HEME AXIAL LIGAND)

(POTENTIAL).

CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).

CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).

CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).

CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).

CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).

CARBOHYD 167 167 A -> P (IN REF. 2).

CONFLICT 160 400 A -> P (IN REF. 2).

CONFLICT 400 524 K -> M (IN REF. 2).

CONFLICT 524 531 Y -> N (IN REF. 2).



```

SQ SEQUENCE 716 AA; 81425 MW; 6E3EBFB975C1FEA4 CRC64;

Query Match
Best Local Similarity 0.5%; Score 7; DB 1; Length 716;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGL 23
Db 6 LALVGL 12

RESULT 91
ID YNC2_CAEEL STANDARD; PRT; 737 AA.
AC P34535;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE HYPOTHETICAL 83.6 KDA PROTEIN R05D3.2 IN CHROMOSOME III.
GN R05D3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Kershaw J.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z48621; CAA88546.1;
DR WormPep; R07B1.9; CE01635.
KW Hypothetical protein.
SQ SEQUENCE 770 AA; 84235 MW; 42EA80C594ACBBD8 CRC64;

Query Match
Best Local Similarity 0.5%; Score 7; DB 1; Length 770;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 468 GYALAGS 474
Db 729 GYALAGS 735

RESULT 93
ID RIR1_VACCC STANDARD; PRT; 771 AA.
AC P20503;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (EC 1.17.4.1)
GN I4L.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10249;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RL Virology 179:517-563(1990).
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'DEOXYRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED
CC THIOREDOXIN + H(2)O -> RIBONUCLEOSIDE DIPHOSPHATE + REDUCED
CC THIOREDOXIN.
CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC LARGE CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M35027; AAA48059.1; -  
DR PIR: I42510; WNV293.  
DR InterPro: IPR000788; -  
DR Pfam: PF00317; ribonucleo\_red; 1.  
DR PRINTS; PR01183; RIBORDTASEM1.  
DR PROSITE; PS00089; RIBORED.LARGE; 1.  
KW Oxidoreductase; DNA replication; Early protein.  
SQ SEQUENCE 771 AA; 87753 MW; 6CA07FC58EDC4F9 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 771;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 GTNGISN 756  
DB 264 GTNGISN 270

RESULT 94  
RIRL\_VACCV STANDARD; PRT; 771 AA.  
AC P12848;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (EC 1.17.4.1)  
DE (RIBONUCLEOSIDE REDUCTASE).  
GN I4L.  
OS Vaccinia virus (strain WR).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Orthopoxvirus.  
OX NCBI\_TaxID=10254;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88215015; PubMed=2835495;  
RA Schmitt J.F.C., Stunnenberg H.G.;  
RT "Sequence and transcriptional analysis of the vaccinia virus HindIII  
RT I fragment.";  
RL J. Virol. 62:1889-1897(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88206055; PubMed=3284177;  
RA Tengelsen L.A., Slabaugh M.B., Bibler J.K., Hruby D.E.;  
RT "Nucleotide sequence and molecular genetic analysis of the large  
RT subunit of ribonucleotide reductase encoded by vaccinia virus.";  
RL Virology 164:121-131(1988).  
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.  
CC -1- CATALYTIC ACTIVITY: 2'DEORO-RIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED  
CC THIOREDOXIN + H(2)O = RIBONUCLEOSIDE DIPHOSPHATE + REDUCED  
CC THIOREDOXIN.  
CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.  
CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.  
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE  
CC LARGE CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: J03399; AAB59806.1; -  
DR EMBL: M20299; AAA48274.1; -

DR PIR: A28611; WZVZH4.  
DR InterPro: IPR000788; -  
DR Pfam: PF00317; ribonucleo\_red; 1.  
DR PRINTS; PR01183; RIBORDTASEM1.  
DR PROSITE; PS00089; RIBORED.LARGE; 1.  
KW Oxidoreductase; DNA replication; Early protein.  
FT CONFLICT 447 447 A -> G (IN REF. 2).  
FT CONFLICT 467 467 A -> V (IN REF. 2).  
FT CONFLICT 526 526 L -> I (IN REF. 2).  
FT CONFLICT 601 606 MPTAST -> LPLDQH (IN REF. 2).  
SQ SEQUENCE 771 AA; 87737 MW; 6CA07FC58F6A8F9 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 771;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 GTNGISN 756  
DB 264 GTNGISN 270

RESULT 95  
KEMK\_MOUSE STANDARD; PRT; 774 AA.  
AC Q05512;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PUTATIVE SERINE/THREONINE-PROTEIN KINASE EMK (EC 2.7.1.-).  
GN EMK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=93364122; PubMed=8358177;  
RA Inglis J.D., Lee M., Hill R.E.;  
RT "Emk, a protein kinase with homologs in yeast maps to mouse  
RT chromosome 19.";  
RL Mamm. Genome 4:401-403(1993).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC NIMI SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X70764; CAA50040.1; -  
DR HSSP: Q63450; IAO6  
DR MGD; MGI:99838; Emk.  
DR InterPro: IPR000719; -  
DR InterPro: IPR002290; -  
DR Pfam: PF00069; kinase; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00108; PROTEIN KINASE-ST; 1.  
DR PROSITE; PS00011; PROTEIN KINASE-DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
FT DOMAIN 53 304 PROTEIN KINASE.  
FT NP\_BIND 59 67 ATP (BY SIMILARITY).  
FT BINDING 82 82 ATP (BY SIMILARITY).  
FT ACT\_SITE 175 175 BY SIMILARITY.  
SQ SEQUENCE 774 AA; 85874 MW; 02BF8D7BF443483A CRC64;

Query Match 0.5%; Score 7; DB 1; Length 774;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 212 GRKASST 218
    |||||
Db 439 GRKASST 445

RESULT 96
RIRL_PLAFG
ID RIRL_PLAFG STANDARD; PRT; 804 AA.
AC P50647;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE (RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (EC 1.17.4.1)
GN RN1.
OS Plasmodium falciparum (isolate PCR-3 / Gambia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5838;
RN [1]
RX MEDLINE=94022359; PubMed=8415692;
RA Rubin H., Salem J.S., Li L.S., Yang F.D., Mama S., Wang Z.M.,
RA Fisher A., Hamann C.S., Cooperman B.S.;
RT "Cloning, sequence determination, and regulation of the
RT ribonucleotide reductase subunits from Plasmodium falciparum: a
RT target for antimalarial therapy.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9280-9284(1993).
CC -|- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -|- CATALYTIC ACTIVITY: 2'DEoxyRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED
CC THIOREDOXIN + H(2)O = RIBONUCLEOSIDE DIPHOSPHATE + REDUCED
CC THIOREDOXIN.
CC -|- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -|- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
CC -|- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC LARGE CHAIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DB EMBL: AF205580; AAA29755.1; -.
DR InterPro: IPR000788; -.
DR Pfam: PF00317; ribonucleo_red; 1.
DR PRINTS: PR01183; RIBORDTASEM1.
DR PROSITE: PS00089; RIBORED_LARGE; 1.
KW Oxidoreductase; DNA replication
SQ SEQUENCE 804 AA; 92402 MW; 8D3C70EA2ED0A6E9 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 804;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 750 GTNGISN 756
    |||||
Db 263 GTNGISN 269

RESULT 97
E2F_DROME
ID E2F_DROME STANDARD; PRT; 805 AA.
AC Q27368;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TRANSCRIPTION FACTOR E2F.
GN E2F OR E2F1.

```

```

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RX SEQUENCE FROM N.A.
RC TISSUE=Eye imaginal disk;
RX MEDLINE=94294381; PubMed=8022787;
RA Dynlacht B.D., Brook A., Dembski M., Yenush L., Dyson N.;
RT "DNA-binding and trans-activation properties of Drosophila E2F and DP
RT proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6359-6363(1994).
RN [2]
RX SEQUENCE FROM N.A.
RC MEDLINE=94158833; PubMed=8114698;
RA Ohtani K., Nevins J.R.;
RT "Functional properties of a Drosophila homolog of the E2F1 gene.";
RL Mol. Cell. Biol. 14:1603-1612(1994).
CC -|- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO E2F SITES.
CC -|- SUBUNIT: HETERODIMER OF E2F AND DP. COOPERATE TO GIVE SEQUENCE-
CC SPECIFIC DNA BINDING AND OPTIMAL TRANS-ACTIVATION.
CC -|- SUBCELLULAR LOCATION: NUCLEAR.
CC -|- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DB EMBL: X78421; CRA55186.1; -.
DR EMBL: U10184; AAA19003.1; -.
DR FlyBase; FBgn0011766; E2f.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 253 318
FT DOMAIN 318 411 DIMERIZATION (POTENTIAL).
FT DOMAIN 14 19 POLY-SER.
FT DOMAIN 64 68 POLY-ASN.
FT DOMAIN 115 125 POLY-ALA.
FT DOMAIN 245 249 POLY-SER.
FT DOMAIN 525 533 POLY-GLN.
FT DOMAIN 594 601 POLY-ALA.
FT DOMAIN 701 710 POLY-GLY.
SQ SEQUENCE 805 AA; 87451 MW; AD652449DDE823B4 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 805;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 948 TATTTLN 954
    |||||
Db 606 TATTTLN 612

RESULT 98
RIRL_PLAF4
ID RIRL_PLAF4 STANDARD; PRT; 806 AA.
AC P50648;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE (RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (EC 1.17.4.1)
GN RN1.
OS Plasmodium falciparum (isolate Dd2).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57267;
RN [1]
RX SEQUENCE FROM N.A.

```

RA MEDLINE-94089761; PubMed-8265664;  
RA Chakrabarti D., Schuster S.M., Chakrabarti R.;  
RT "Cloning and characterization of subunit genes of ribonucleotide  
RT reductase, a cell-cycle-regulated enzyme, from Plasmodium  
RT falciparum.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:12020-12024(1993).  
CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.  
CC -!- CATALYTIC ACTIVITY: 2'DEOXYRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED  
CC THIOREDOXIN + H(2)O = RIBONUCLEOSIDE DIPHOSPHATE + REDUCED  
CC THIOREDOXIN.  
CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.  
CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.  
CC -!- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE  
CC LARGE CHAIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U01323; AAA50171.1; -  
DR HSSP; P00452; 4R1R.  
DR InterPro; IPR000788; -  
DR Pfam; PF00317; ribonucleo\_red; 1.  
DR PRINTS; P01183; RIBORDTASEM1.  
DR PROSITE; PS00089; RIBORED\_LARGE; 1.  
KW Oxidoreductase; DNA replication.  
SQ SEQUENCE 806 AA; 92512 MW; A23B414A687B275F CRC64;  
  
Query Match 0.5%; Score 7; DB 1; Length 806;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 750 GTNGISN 756  
DB 262 GTNGISN 268  
|||||||  
  
RESULT 99  
DYNL CAEEL STANDARD; PRT; 830 AA.  
AC P39055;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE DYNAMIN.  
GN DYN-1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-BRISTOL N2;  
RC MEDLINE-97439883; PubMed-9294229;  
RX Clark S.G., Shurland D.L., Meyerowitz E.M., Bargmann C.I.,  
RA van der Blik A.M.;  
RA "A dynamin GTPase mutation causes a rapid and reversible temperature-  
RT inducible locomotion defect in C. elegans.";  
RT Proc. Natl. Acad. Sci. U.S.A. 94:10438-10443(1997).  
RL [2]  
RN REVISIONS TO C-TERMINUS.  
RP STRAIN-BRISTOL N2;  
RC van der Blik A.M.;  
RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED  
CC IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE  
CC GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN  
CC PARTICULAR ENDOCYTOSIS.

CC -!- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.  
CC -!- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L29031; AAB72228.2; -  
DR HSSP; Q05193; 2DYN.  
DR InterPro; IPR000375; -  
DR InterPro; IPR001401; -  
DR InterPro; IPR001849; -  
DR Pfam; PF00169; PH; 1.  
DR Pfam; PF00350; dynamin; 1.  
DR Pfam; PF01031; dynamin\_2; 1.  
DR PRINTS; PR00195; DYNAMIN.  
DR PROSITE; PS00410; DYNAMIN; 1.  
DR PROSITE; PS00003; PH\_DOMAIN; 1.  
KW Motor protein; GTP-binding; Microtubules; Multigene family;  
KW Endocytosis.  
FT NP\_BIND 40 47 GTP (BY SIMILARITY).  
FT NP\_BIND 138 142 GTP (BY SIMILARITY).  
FT NP\_BIND 207 210 GTP (BY SIMILARITY).  
FT DOMAIN 519 624 PH.  
SQ SEQUENCE 830 AA; 93348 MW; FC3D7106D079EDC5 CRC64;  
  
Query Match 0.5%; Score 7; DB 1; Length 830;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1261 LKLAKEV 1267  
DB 189 LKLAKEV 195  
|||||||  
  
RESULT 100  
PSP1\_YEAST STANDARD; PRT; 841 AA.  
AC P50896;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PSP1 PROTEIN (GIN5 PROTEIN).  
GN PSP1 OR GIN5 OR YDR505C OR D9719.11.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,  
RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,  
RA Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,  
RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,  
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,  
RA Winant A., Yellon M., Botstein D., Davis R.W.;  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-98190324; PubMed-9529527;  
RA Formosa T., Nittis T.;  
RT "Suppressors of the temperature sensitivity of DNA polymerase alpha  
RT mutations in Saccharomyces cerevisiae.";  
RL Mol. Gen. Genet. 257:461-468(1998).  
CC -!- FUNCTION: DNA POLYMERASE ALPHA MUTATION SUPPRESSOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC use the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: U33057; AAB64947.1; -  
DR EMBL: U33115; AAA93076.1; -  
DR SGD: S0002913; PSPI.  
FT DOMAIN 98 101 POLY-GLN.  
FT DOMAIN 139 144 POLY-ASN.  
FT DOMAIN 314 320 POLY-ASN.  
FT DOMAIN 348 359 POLY-GLN.  
FT CONFLICT 116 116 S -> C (IN REF. 2).  
FT CONFLICT 122 133 QMSSNSEPMISA -> KCLRILQSCVP (IN REF. 2).  
FT CONFLICT 198 210 GSNFAAPSHSAGN -> RAILPHHTVLAT (IN  
REF. 2).  
FT CONFLICT 732 732 E -> K (IN REF. 2).  
FT SEQUENCE 841 AA; 95348 MW; 53B5C2753C5A7FF7 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 841;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 QFNGNSF 175  
Db 160 QFNGNSF 166  
|||||

RESULT 101  
DYN3\_RAT  
ID DYN3\_RAT STANDARD; PRT; 848 AA.  
AC Q08877;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE DYNAMIN 3 (DYNAMIN, TESTICULAR) (T-DYNAMIN).  
GN DNM3 OR DYN3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Testis;  
RX MEDLINE-93366923; PubMed-8360266;  
RA Nakata T., Takamura R., Hirokawa N.;  
RT "A novel member of the dynamin family of GTP-binding proteins is  
expressed specifically in the testis.";  
RL J. Cell Sci. 105:1-5(1993).  
CC -1- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED  
IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE  
GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN  
PARTICULAR ENDOCYTOSIS.  
CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN GERM-CELL-DEPLETED TESTIS,  
INDICATING ITS EXPRESSION IN SERTOLI CELLS.  
CC -1- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D14076; BAA03161.1; -  
DR HSSP: Q05193; 2DYN.  
DR InterPro: IPR000375; -

DR InterPro: IPR001401; -  
DR InterPro: IPR001849; -  
DR Pfam: PF00169; PH; 1.  
DR Pfam: PF00350; dynamin; 1.  
DR Pfam: PF01031; dynamin\_2; 1.  
DR PRINTS: PR00195; DYNAMIN.  
DR PROSITE: PS00410; DYNAMIN; 1.  
DR PROSITE: PS00003; PH DOMAIN; 1.  
KW Motor protein; GTP-binding; Microtubules; Multigene family;  
KW Endocytosis.  
FT NP\_BIND 38 45 GTP (BY SIMILARITY).  
FT NP\_BIND 136 140 GTP (BY SIMILARITY).  
FT NP\_BIND 205 208 GTP (BY SIMILARITY).  
FT DOMAIN 515 621 PH.  
SQ SEQUENCE 848 AA; 95595 MW; 802E365FCFC685F6 CRC64;  
  
Query Match 0.5%; Score 7; DB 1; Length 848;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1261 LKLAKEV 1267  
Db 187 LKLAKEV 193  
|||||

RESULT 102  
DYN2\_MOUSE  
ID DYN2\_MOUSE STANDARD; PRT; 866 AA.  
AC P39054;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE DYNAMIN 2 (DYNAMIN UDNM).  
GN DNM2 OR DYN2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-NIH SWISS;  
RX MEDLINE-97288532; PubMed-9143510;  
RA Klocke R., Augustin A., Ronsiek M., Stief A., van der Putten H.,  
RA Jockusch H.;  
RT "Dynammin genes Dnm1 and Dnm2 are located on proximal mouse  
chromosomes 2 and 9, respectively.";  
RL Genomics 41:290-292(1997).  
CC -1- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED  
IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE  
GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN  
PARTICULAR ENDOCYTOSIS.  
CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.  
CC -1- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L31398; AAA40523.1; -  
DR HSSP: Q05193; 1DYN.  
DR MGD: MGI:109547; Dnm2.  
DR InterPro: IPR000375; -  
DR InterPro: IPR001401; -  
DR InterPro: IPR001849; -  
DR Pfam: PF00169; PH; 1.  
DR Pfam: PF00350; dynamin; 1.  
DR Pfam: PF01031; dynamin\_2; 1.

DR PRINTS; PRO0195; DYNAMIN.  
DR PROSITE; PS00410; DYNAMIN; 1.  
DR PROSITE; PS50003; PH\_DOMAIN; 1.  
KW Motor protein; GTP-binding; Microtubules; Multigene family;  
KW Endocytosis.  
FT NP\_BIND 38 45 GTP (BY SIMILARITY).  
FT NP\_BIND 136 140 GTP (BY SIMILARITY).  
FT NP\_BIND 205 208 GTP (BY SIMILARITY).  
FT DOMAIN 515 621 PH.  
SQ SEQUENCE 866 AA; 97683 MW; CC52BBBC26B5D757 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 866;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1261 LKLAKEV 1267  
Db 187 LKLAKEV 193

RESULT 103  
DYN2\_HUMAN STANDARD; PRT; 870 AA.  
AC P50570;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE DYNAMIN 2.  
GN DNM2 OR DYN2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96011652; PubMed=7590285;  
RA Diatloff-Zito C., Gordon A.J.E., Duchaud E., Merlin G.;  
RT "Isolation of an ubiquitously expressed cDNA encoding human dynamin II, a member of the large GTP-binding protein family.";  
RL Gene 163:301-306(1995).  
CC -!- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN PARTICULAR ENDOCYTOSIS.  
CC -!- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.  
CC -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.  
CC -!- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L36983; AAA88025.1;  
CC DR HSSP; Q05193; 2DYN.  
CC MM; 602378;  
CC InterPro; IPR000375;  
CC InterPro; IPR001401;  
CC InterPro; IPR001849;  
CC Pfam; PF00169; PH; 1.  
CC Pfam; PF00350; dynamin; 1.  
CC Pfam; PF01031; dynamin\_2; 1.  
CC PRINTS; PRO0195; DYNAMIN.  
CC PROSITE; PS00410; DYNAMIN; 1.  
CC PROSITE; PS50003; PH\_DOMAIN; 1.  
KW Motor protein; GTP-binding; Microtubules; Multigene family;  
KW Endocytosis; Alternative splicing.  
FT NP\_BIND 38 45 GTP (BY SIMILARITY).

FT NP\_BIND 136 140 GTP (BY SIMILARITY).  
FT NP\_BIND 205 208 GTP (BY SIMILARITY).  
FT DOMAIN 519 625 PH.  
FT VARSPLIC 516 519 MISSING (IN A ISOFORM).  
SQ SEQUENCE 870 AA; 98018 MW; 149189A598BE7039 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 870;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1261 LKLAKEV 1267  
Db 187 LKLAKEV 193

RESULT 104  
DYN2\_RAT STANDARD; PRT; 870 AA.  
AC P39052;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE DYNAMIN 2.  
GN DNM2 OR DYN2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RX MEDLINE=94119943; PubMed=8290576;  
RA Cook T.A., Urrutia R., McNiven M.A.;  
RT "Identification of dynamin 2, an isoform ubiquitously expressed in rat tissues.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:644-648(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=94140890; PubMed=8308025;  
RA Sontag J.-M., Fykse E.M., Ushkaryov Y., Liu J.-P., Robinson P.J., Suedhof T.C.;  
RT "Differential expression and regulation of multiple dynamins.";  
RL J. Biol. Chem. 269:4547-4554(1994).  
CC -!- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN PARTICULAR ENDOCYTOSIS.  
CC -!- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.  
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; IIAA, IIBA (SHOWN HERE) AND IIC; ARE PRODUCED BY ALTERNATIVE SPLICING. IIBA DIFFERS FROM IIAA BY THE REPLACEMENT OF AN EXON BY ANOTHER ONE WHICH IS HIGHLY SIMILAR.  
CC -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, INCLUDING THE BRAIN.  
CC HIGHEST LEVELS IN THE TESTIS.  
CC -!- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L25605; AAA19736.1;  
CC DR EMBL; L24562; AAA16745.1;  
CC HSSP; Q05193; 2DYN.  
CC InterPro; IPR000375;  
CC InterPro; IPR001401;  
CC InterPro; IPR001849;  
CC Pfam; PF00169; PH; 1.

```
DR Pfam: PF00350; dynamin; 1.
DR Pfam: PF01031; dynamin_2; 1.
DR PRINTS; PRO0195; DYNAMIN.
DR PROSITE; PS00410; DYNAMIN; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR Motor protein; GTP-binding; Microtubules; Multigene family;
KW Endocytosis; Alternative splicing.
FT NP_BIND 38 45 GTP (BY SIMILARITY).
FT NP_BIND 136 140 GTP (BY SIMILARITY).
FT NP_BIND 205 208 GTP (BY SIMILARITY).
FT DOMAIN 519 625 PH.
FT VARSPPLIC 407 444 LAPEAIVKQVKKLPCPKCLKVDLVQIELISTVROCT'S ->
MAPEAIVKQVKKLPCPKCLKVDLVQIELISTVROCT'S ->
(IN ISOFORM IIAA).
FT VARSPPLIC 516 519 MISSING (IN ISOFORM IIC).
FT CONFLICT 298 298 S -> T (IN REF. 1).
FT CONFLICT 389 389 S -> T (IN REF. 1).
FT CONFLICT 487 487 N -> K (IN REF. 1).
FT CONFLICT 637 637 G -> E (IN REF. 1).
FT CONFLICT 719 719 MISSING (IN REF. 1).
FT CONFLICT 786 791 GPTPGP -> PHTGA (IN REF. 1).
SQ SEQUENCE 870 AA; 98230 MW; 3802DBAFA3ABBE98 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 870;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1261 LKLAKEV 1267
|||||||
DB 187 LKLAKEV 193

-RESULT 105
DYN_DROME
ID DYN_DROME STANDARD; PRT; 883 AA.
AC P27619;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE DYNAMIN (SHIBIRE PROTEIN).
GN SHI.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RX MEDLINE=91260878; PubMed=1828536;
RA Chen M.S., Obar R.A., Schroeder C.C., Austin T.W., Poody C.A.,
RA Wadsworth S.C., Vallee R.B.;
RT "Multiple forms of dynamin are encoded by shibire, a Drosophila gene
involved in endocytosis.";
RL Nature 351:583-586(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=91238973; PubMed=1674590;
RA van der Bliek A.M., Meyerowitz E.M.;
RT "Dynamin-like protein encoded by the Drosophila shibire gene
associated with vesicular traffic.";
RL Nature 351:411-414(1991).
CC -1- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN WHICH IS
INVOLVED IN THE PRODUCTION OF MICROTUBULE BUNDLES AND WHICH IS
ABLE TO BIND AND HYDROLYZE GTP. SHIBIRE IS IMPLICATED IN ENDOCYTIC
PROTEIN SORTING.
CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST THREE ISOFORMS ARE CREATED BY
ALTERNATIVE SPLICING OF THE SHI GENE.
CC -1- DISEASE: SHIBIRE MUTATION IS THE CAUSE OF TEMPERATURE-SENSITIVE
PARALYSIS. THIS IS BELIEVED TO BE DUE TO A REVERSIBLE BLOCK OF
```

```
CC CC
CC ENDOCYTOSIS, WHICH PREVENTS MEMBRANE CYCLING AND THUS DEPLETES
SYNAPTIC VESICLES.
CC -1- MISCELLANEOUS: SHIBIRE IS A JAPANESE WORD THAT MEANS "PARALYZED".
CC -1- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X59449; CAA42068.1; -
DR EMBL; X59448; CAA42067.1; -
DR EMBL; X59435; CAA42061.1; -
DR EMBL; X59436; CAA42062.1; -
DR PIR; S15413; S15413.
DR PIR; S15497; S15497.
DR PIR; S15498; S15498.
DR PIR; S16130; S16130.
DR PIR; S17974; S17974.
DR PIR; S17975; S17975.
DR HSP; Q05193; 2DXN.
DR Flybase; FBgn0003392; shi.
DR InterPro; IPR000375; -
DR InterPro; IPR001401; -
DR InterPro; IPR001849; -
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00350; dynamin; 1.
DR Pfam; PF01031; dynamin_2; 1.
DR PRINTS; PRO0195; DYNAMIN.
DR PROSITE; PS00410; DYNAMIN; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Motor protein; GTP-binding; Microtubules; Alternative splicing;
KW Endocytosis.
FT NP_BIND 33 40 GTP (POTENTIAL).
FT NP_BIND 131 135 GTP (POTENTIAL).
FT NP_BIND 200 203 GTP (POTENTIAL).
FT DOMAIN 513 621 PH.
FT DOMAIN 750 833 PRO-RICH.
FT VARSPPLIC 635 640 MISSING (IN THIRD ISOFORM).
FT VARSPPLIC 836 836 V -> R (IN SHORT ISOFORM).
FT VARSPPLIC 837 883 MISSING (IN SHORT ISOFORM).
FT VARIANT 141 141 G -> S (IN SHI-TS2 MUTANT).
FT VARIANT 268 268 G -> D (IN SHI-TS1 MUTANT).
FT VARIANT 594 594 R -> K.
SQ SEQUENCE 883 AA; 98537 MW; F428C7AA708C7C70 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 883;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1261 LKLAKEV 1267
|||||||
DB 182 LKLAKEV 188

RESULT 106
NUOG_BUCAI
ID NUOG_BUCAI STANDARD; PRT; 906 AA.
AC P57257;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NADH DEHYDROGENASE I CHAIN G (EC 1.6.5.3) (NADH-UBIQUINONE
OXIDOREDUCTASE CHAIN G).
GN NUOG OR BUI59.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
```

NCBI\_TaxID=118099;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TOKYO 1998;  
 RA MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 Buchnera sp. APS.";  
 RL Nature 407:81-86(2000).  
 CC -I- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -I- COFACTOR: MAY BIND TWO 4FE-4S CLUSTER AND ONE 2FE-2S CLUSTER.  
 CC -I- SUBUNIT: COMPOSED OF 13 DIFFERENT SUBUNITS. SUBUNITS NUOCD, E, F,  
 CC AND G CONSTITUTE THE PERIPHERAL SECTOR OF THE COMPLEX (BY  
 CC SIMILARITY).  
 CC -I- SIMILARITY: BELONGS TO THE COMPLEX I 75 KDA SUBUNIT FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AP001118; BAB12877.1; .  
 DR PROSITE: PS00641; COMPLEX1\_75K\_1; 1.  
 DR PROSITE: PS00642; COMPLEX1\_75K\_2; 1.  
 DR PROSITE: PS00643; COMPLEX1\_75K\_3; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Iron-sulfur; 4Fe-4S.  
 FT METAL 23 23 IRON-SULFUR (2FE-2S) (POTENTIAL).  
 FT METAL 34 34 IRON-SULFUR (2FE-2S) (POTENTIAL).  
 FT METAL 45 45 IRON-SULFUR (2FE-2S) (POTENTIAL).  
 FT METAL 48 48 IRON-SULFUR (2FE-2S) (POTENTIAL).  
 FT METAL 103 103 IRON-SULFUR (2FE-2S) (POTENTIAL).  
 FT METAL 106 106 IRON-SULFUR (2FE-2S) (POTENTIAL).  
 FT METAL 112 112 IRON-SULFUR (2FE-2S) (POTENTIAL).  
 FT METAL 151 151 IRON-SULFUR (4FE-4S) (POTENTIAL).  
 FT METAL 154 154 IRON-SULFUR (4FE-4S) (POTENTIAL).  
 FT METAL 157 157 IRON-SULFUR (4FE-4S) (POTENTIAL).  
 FT METAL 201 201 IRON-SULFUR (4FE-4S) (POTENTIAL).  
 SQ SEQUENCE 906 AA; 103761 MW; A1CBA1210C087A54 CRC64;  
  
 Query Match 0.5%; Score 7; DB 1; Length 906;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 97 EYDLYRS 103  
 I I I I I I I  
 Db 571 EYDLYRS 577  
  
 RESULT 107  
 MSH2.HUMAN  
 ID MSH2.HUMAN STANDARD; PRT; 934 AA.  
 AC P43246;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE DNA MISMATCH REPAIR PROTEIN MSH2.  
 GN MSH2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94073959; PubMed=8252616;  
 RA Fishel R., Lescoe M., Rao M., Copeland N.G., Jenkins N.A.,  
 RA Garber J., Kane M., Kolodner R.D.;  
 RT "The human mutator gene homolog MSH2 and its association with  
 RT hereditary nonpolyposis colon cancer.";  
 RL Cell 75:1027-1038(1993).

RN  
 RP ERATUM.  
 RX MEDLINE=94208055; PubMed=8156592;  
 RA Fishel R., Lescoe M., Rao M., Copeland N.G., Jenkins N.A.,  
 RA Garber J., Kane M., Kolodner R.D.;  
 RL Cell 77:167-167(1994).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=95229152; PubMed=7713503;  
 RA Kolodner R.D., Hall N.R., Lipford J., Kane M.F., Rao M.R.,  
 RA Morrison P., Wirth L., Finan P.J., Burn J., Chapman P., Earabino C.,  
 RA Merchant E., Bishop D.T.;  
 RT "Structure of the human MSH2 locus and analysis of two Muir-Torrie  
 RT kindreds for msh2 mutations.";  
 RL Genomics 24:516-526(1994).  
 [4]  
 RN DNA-BINDING.  
 RX MEDLINE=95007585; PubMed=7923193;  
 RA Fishel R., Ewel A., Lescoe M.K.;  
 RT "Purified human MSH2 protein binds to DNA containing mismatched  
 RT nucleotides.";  
 RL Cancer Res. 54:5539-5542(1994).  
 [5]  
 RN REVIEW.  
 RP MEDLINE=94310688; PubMed=8036718;  
 RA Jiricny J.;  
 RT "Colon cancer and DNA repair: have mismatches met their match?";  
 RL Trends Genet. 10:164-168(1994).  
 [6]  
 RN REVIEW ON VARIANTS.  
 RP MEDLINE=97403931; PubMed=9259192;  
 RA Papadopoulos N., Lindblom A.;  
 RT "Molecular basis of HNPCC: mutations of MMR genes.";  
 RL Hum. Mutat. 10:89-99(1997).  
 [7]  
 RN VARIANTS HNPCC LEU-622 AND TYR-639.  
 RX MEDLINE=94084796; PubMed=8261515;  
 RA Leach F.S., Nicolaides N.C., Papadopoulos N., Liu B., Jen J.,  
 RA Parsons R., Peltomaki P., Sistonen P., Aaltonen L.A., Yu J.-W.,  
 RA Nystrom-Lahti M., Guan X.-Y., Zhang J., Meltzer P.S., Todd S.,  
 RA Rao F.T., Chen D.J., Cerosaletti K.M., Fournier R.E.K., Todd S.,  
 RA Lewis T., Leach R.J., Naylor S.L., Weissbach J., Mecklin J.-P.,  
 RA Jaarvinen H., Petersen G.M., Hamilton S.R., Green J., Jass J.,  
 RA Watson P., Lynch H.T., Trent J.M., de la Chapelle A., Kinzler K.W.,  
 RA Vogelstein B.;  
 RT "Mutations of a mutS homolog in hereditary nonpolyposis colorectal  
 RT cancer.";  
 RL Cell 75:1215-1225(1993).  
 [8]  
 RN VARIANT HNPCC ASN-596 DEL.  
 RP MEDLINE=95179130; PubMed=7874129;  
 RA Mary J.-L., Bishop T., Kolodner R., Lipford J.R., Kane M., Weber W.,  
 RA Thorhorst J., Mueller H., Spycher M., Scott R.J.;  
 RT "Mutational analysis of the hMSH2 gene reveals a three base pair  
 RT deletion in a family predisposed to colorectal cancer development.";  
 RL Hum. Mol. Genet. 3:2067-2069(1994).  
 [9]  
 RN VARIANT HIS-96.  
 RP MEDLINE=95243220; PubMed=7726159;  
 RX Wijnen J., Vasen H., Khan P.M., Menko F.H., van der Klift H.,  
 RA van Leeuwen C., van den Broek M., van Leeuwen-Cornelisse I.,  
 RA Nagengast F., Meijers-Heijboer A., Lindhout D., Griffioen G., Cats A.,  
 RA Kleibouker J., Varesco L., Bertario L., Bisgaard M.L., Mohr J.,  
 RA Fodde R.;  
 RT "Seven new mutations in hMSH2, an HNPCC gene, identified by  
 RT denaturing gradient-gel electrophoresis.";  
 RL Am. J. Hum. Genet. 56:1060-1066(1995).  
 [10]  
 RN VARIANT HNPCC ASP-322.  
 RP MEDLINE=96163505; PubMed=8566964;  
 RX Maliaka Y.K., Chudina A.P., Belev N.F., Alday P., Bochkov N.P.,  
 RA Buerstedde J.-M.;  
 RT "CpG dinucleotides in the hMSH2 and hMLH1 genes are hotspots for



HNPPC mutations.";  
 Hum. Genet. 97:251-255(1996).  
 [11]  
 VARIANT GLN-46.  
 MEDLINE-96293410; PubMed-8700523;  
 RA Bubb V.J., Curtis I.J., Cunningham C., Dunlop M.G., Carothers A.D.,  
 RA Morris R.G., White S., Bird C.C., Wylie A.H.;  
 RT "Microsatellite instability and the role of hMSH2 in sporadic  
 RT colorectal cancer.";  
 RL Oncogene 12:2641-2649(1996).  
 RN [12]  
 VARIANTS HNPPC THR-305; THR-834 AND ASN-596 DEL.  
 MEDLINE-97456423; PubMed-9311737;  
 RA Wijnen J., Khan P.M., Vasen H., van der Klift H., Mulder A.,  
 RA van Leeuwen-Cornelisse I., Bakker B., Losekoot M., Moeller P.,  
 RA Podde R.;  
 RT "Hereditary nonpolyposis colorectal cancer families not complying  
 RT with the Amsterdam criteria show extremely low frequency of  
 RT mismatch-repair-gene mutations.";  
 RL Am. J. Hum. Genet. 61:329-335(1997).  
 RN [13]  
 VARIANT HNPPC VAL-562.  
 MEDLINE-97201114; PubMed-9048925;  
 RA Beck N.E., Tomlinson I.P.M., Homfray T., Frayling I., Hodgson S.V.,  
 RA Harocopos C., Bodmer W.F.;  
 RT "Use of SSCP analysis to identify germline mutations in HNPPC  
 RT families fulfilling the Amsterdam criteria.";  
 RL Hum. Genet. 99:219-224(1997).  
 RN [14]  
 VARIANTS HNPPC ASP-322 AND PHE-697.  
 MEDLINE-97442278; PubMed-9298827;  
 RA Wehner M., Buschhausen L., Lamberti C., Kruse R., Caspari R.,  
 RA Propping P., Friedl W.;  
 RT "Hereditary nonpolyposis colorectal cancer (HNPPC): eight novel  
 RT germline mutations in hMSH2 or hMLH1 genes.";  
 RL Hum. Mutat. 10:241-244(1997).  
 RN [15]  
 VARIANT HNPPC VAL-265--GLN-314 DEL, AND VARIANTS GLY-641 AND VAL-770.  
 MEDLINE-98386069; PubMed-9718327;  
 RA Farrington S.M., Lin-Goerke J., Ling J., Wang Y., Burczak J.D.,  
 RA Robbins D.J., Dunlop M.G.;  
 RT "Systematic analysis of hMSH2 and hMLH1 in young colon cancer patients  
 RT and controls.";  
 RL Am. J. Hum. Genet. 63:749-759(1998).  
 RN [16]  
 VARIANT PHE-390.  
 MEDLINE-98284542; PubMed-9621522;  
 RA Okamura S., Koyama K., Miyoshi Y., Monden M., Takami M.;  
 RT "Novel germline mutations of hMSH2 in a patient with hereditary  
 RT nonpolyposis colorectal cancer 'HNPPC' and in a patient with six  
 RT primary cancers.";  
 RL J. Hum. Genet. 43:143-145(1998).  
 RN [17]  
 VARIANTS HNPPC ARG-692 AND ARG-697.  
 MEDLINE-20081064; PubMed-10612836;  
 RA Isidro G., Veiga I., Matos P., Almeida S., Bizarro S., Marshall B.,  
 RA Baptista M., Leite J., Regateiro F., Soares J., Castedo S.,  
 RA Boavida M.G.;  
 RT "Four novel MSH2 / MLH1 gene mutations in portuguese HNPPC families.";  
 RL Hum. Mutat. 15:116-116(2000).  
 CC -1- FUNCTION: INVOLVED IN POSTREPLICATION MISMATCH REPAIR. BINDS  
 CC SPECIFICALLY TO DNA CONTAINING MISMATCHED NUCLEOTIDES THUS  
 CC PROVIDING A TARGET FOR THE EXCISION REPAIR PROCESSES  
 CC CHARACTERISTIC OF POSTREPLICATION MISMATCH REPAIR.  
 CC  
 CC -1- SUBUNIT: HETERODIMER OF MSH2 AND MSH6 (GTBP).  
 CC  
 CC -1- DISEASE: DEFECTS IN MSH2 ARE A CAUSE OF HEREDITARY NONPOLYPOSIS  
 CC COLORECTAL CANCER (HNPPC) (LYNCH SYNDROME). HNPPC IS AN AUTOSOMAL,  
 CC DOMINANTLY INHERITED DISEASE ASSOCIATED WITH MARKED INCREASE IN  
 CC CANCER SUSCEPTIBILITY. IT IS CHARACTERIZED BY A FAMILIAL  
 CC PREDISPOSITION TO EARLY ONSET COLORECTAL CARCINOMA AND EXTRA-  
 CC COLORECTAL CANCERS OF THE GASTROINTESTINAL, UROLOGICAL AND FEMALE  
 CC REPRODUCTIVE TRACTS. HNPPC IS REPORTED TO BE THE MOST COMMON FORM  
 CC OF INHERITED COLORECTAL CANCER, ACCOUNTING FOR ABOUT 1% OR MORE OF

ALL COLORECTAL CANCER CASES. DEFECTS IN TWO OF THE KNOWN MISMATCH  
 REPAIR GENES (NAMELY MSH2 AND MSH1) ACCOUNT FOR OVER 90% OF  
 MUTATIONS FOUND IN HNPPC FAMILIES. CANCERS IN HNPPC ORIGINATE  
 WITHIN BENIGN NEOPLASTIC POLYPS TERMED ADENOMAS.  
 -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; L47583; AAB59564.1; -  
 EMBL; L47582; AAB59565.1; -  
 EMBL; L47581; AAA76858.1; -  
 EMBL; U04045; AAA61870.1; -  
 EMBL; U03911; AAA18643.1; -  
 EMBL; U41221; AAA82080.1; ALT-SEQ.  
 EMBL; U41206; AAA82080.1; JOINED.  
 EMBL; U41207; AAA82080.1; JOINED.  
 EMBL; U41208; AAA82080.1; JOINED.  
 EMBL; U41210; AAA82080.1; JOINED.  
 EMBL; U41211; AAA82080.1; JOINED.  
 EMBL; U41212; AAA82080.1; JOINED.  
 EMBL; U41213; AAA82080.1; JOINED.  
 EMBL; U41214; AAA82080.1; JOINED.  
 EMBL; U41215; AAA82080.1; JOINED.  
 EMBL; U41216; AAA82080.1; JOINED.  
 EMBL; U41217; AAA82080.1; JOINED.  
 EMBL; U41218; AAA82080.1; JOINED.  
 EMBL; U41219; AAA82080.1; JOINED.  
 EMBL; U41220; AAA82080.1; JOINED.  
 MIM: 120435; -

Query Match 0.5%; Score 7; DB 1; Length 934;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1010 LESAAEV 1016  
 Db 11 LESAAEV 17  
 |||||  
 RESULT 108  
 V120\_HSV7J  
 ID V120\_HSV7J STANDARD; PRT; 938 AA.  
 AC P52438;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CAPSID ASSEMBLY PROTEIN U30.  
 GN U30.  
 OS Human herpesvirus (type 7 / strain JT) (HHV7).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 OX NCBI\_TaxID=57278;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nicholas J.;  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL37,  
 CC HSV-1 23, EBV BOLF1, VZV 21, HSV-1 63, AND HCMV UL47.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----



```
DR EMBL; U43400; AAC54692.1; -
KW Capsid assembly.
SQ SEQUENCE 938 AA; 110170 MW; F4E39A2BF0D32BC9 CRC64;

Query Match
Best Local Similarity 0.5%; Score 7; DB 1; Length 938;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1014 AEVLQF 1020
DB 112 AEVLQF 118

RESULT 109.
N120_YEAST
ID N120_YEAST STANDARD; PRT; 1037 AA.
AC P35729; P35730;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NUCLEOPORIN NUP120 (NUCLEAR PORE PROTEIN NUP120).
GN NUP120 OR RAT2 OR YKL057C OR YKL314 OR YKL313.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94378723; PubMed=8091862;
RA Rasmussen S.W.;
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the
RT FBAL and TOA2 genes, an open reading frame (ORF) similar to a
RT translationally controlled tumour protein, one ORF containing motifs
RT also found in plant storage proteins and 13 ORFs with weak or no
RT homology to known proteins.";
RL Yeast 10:S63-S68(1994).
RN [2]
RP REVISIONS.
RA Rasmussen S.W.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION, AND SEQUENCE OF 189-206 AND 800-807.
RX MEDLINE=96134019; PubMed=8557736;
RA Aitchison J.D., Blobel G., Rout M.P.;
RT "Nup120p: a yeast nucleoporin required for NPC distribution and mRNA
RT transport.";
RL J. Cell Biol. 131:1659-1676(1995).
RN [4]
RP PARTIAL SEQUENCE OF 550-555; 799-803 AND 375-384.
RX MEDLINE=96152656; PubMed=8565072;
RA Siniosoglou S., Wimmer C., Rieger M., Doye V., Tekotte H., Weise C.,
RA Emig S., Segref A., Hurt E.C.;
RT "A novel complex of nucleoporins, which includes Sec13p and a Sec13p
RT homolog, is essential for normal nuclear pores.";
RL Cell 84:265-275(1996).
CC -1- FUNCTION: REQUIRED FOR EFFICIENT MRNA EXPORT FROM THE NUCLEUS TO
CC THE CYTOPLASM AND FOR CORRECT NUCLEAR PORE BIOGENESIS.
CC -1- SUBUNIT: INTERACTS WITH NUP84, NUP85, SEC13 AND SEH1.
CC -1- SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z46937; CAA87056.1; -
DR WormPep; F43C1.1; CE01582.
DR InterPro; IPR001611; -
DR InterPro; IPR001932; -
DR Pfam; PF00560; LRR; 13.
DR Pfam; PF00481; PP2C; 1.
DR PRINTS; PR00019; LEURICHRPT.
KW Hypothetical protein; Leucine-rich repeat; Repeat.
FT REPEAT 137 160
FT REPEAT 161 185
FT REPEAT 199 222
FT REPEAT 223 245
FT REPEAT 245 268
FT REPEAT 269 291
FT REPEAT 293 316
FT REPEAT 332 358
FT REPEAT 364 389
FT REPEAT 396 419
FT REPEAT 419 442
FT REPEAT 444 464
FT REPEAT 465 487
FT REPEAT 489 514
FT REPEAT 536 559
FT REPEAT 561 585
FT REPEAT 587 605
FT REPEAT 605 627

Query Match
Best Local Similarity 0.5%; Score 7; DB 1; Length 1037;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1013 AAEVLQ 1019
DB 977 AAEVLQ 983

RESULT 110
YR7L_CAEEL
ID YR7L_CAEEL STANDARD; PRT; 1039 AA.
AC Q09564;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 118.2 KDA PROTEIN F43C1.1 IN CHROMOSOME III.
GN F43C1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Jassal B.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 18 LEUCINE-RICH REPEATS (LRR).
CC -1- SIMILARITY: CONTAINS A PP2C-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z46937; CAA87056.1; -
DR WormPep; F43C1.1; CE01582.
DR InterPro; IPR001611; -
DR InterPro; IPR001932; -
DR Pfam; PF00560; LRR; 13.
DR Pfam; PF00481; PP2C; 1.
DR PRINTS; PR00019; LEURICHRPT.
KW Hypothetical protein; Leucine-rich repeat; Repeat.
FT REPEAT 137 160
FT REPEAT 161 185
FT REPEAT 199 222
FT REPEAT 223 245
FT REPEAT 245 268
FT REPEAT 269 291
FT REPEAT 293 316
FT REPEAT 332 358
FT REPEAT 364 389
FT REPEAT 396 419
FT REPEAT 419 442
FT REPEAT 444 464
FT REPEAT 465 487
FT REPEAT 489 514
FT REPEAT 536 559
FT REPEAT 561 585
FT REPEAT 587 605
FT REPEAT 605 627
```

FT DOMAIN 669 903 PP2C-LIKE.  
SQ SEQUENCE 1039 AA; 118182 MW; 0C6D1AFC79A0D32A CRC64;

Query Match 0.5%; Score 7; DB 1; Length 1039;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 966 LQTLALS 972  
Db 399 LQTLALS 405

RESULT 111  
IRRL\_YEAST STANDARD; PRT; 1150 AA.

AC P40541; 1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE IRL1 PROTEIN  
GN IRL1 OR Y1026C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;  
RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
RA Louis E., Lye G., Moule S., Moutie T., Odell C., Pearson D.,  
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
RA Walsh S.V., Whitehead S.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP CHARACTERIZATION.

RX MEDLINE=96090137; PubMed=7483852;  
RA Kurlandzka A., Rytka J., Gromadka R., Murawski M.;  
RT "A new essential gene located on Saccharomyces cerevisiae chromosome IX.";  
RL Yeast 11:885-890(1995).  
CC -!- FUNCTION: NOT KNOWN; MAY PARTICIPATE IN CELL-CELL AND/OR CELL-SUBSTRATUM INTERACTIONS.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: SOME, TO S.POMBE SPAC17H9.20.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Z46881; CAAB6966.1; -.  
DR SGD; S0001288; IRL1.  
FT DOMAIN 44 48 POLY-GLU.  
FT DOMAIN 65 70 POLY-ASP.  
SQ SEQUENCE 1150 AA; 133008 MW; 89688EA09485AC28 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 1150;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 ESSONNS 364  
Db 1120 ESSONNS 1126

RESULT 112

YOP4\_CAEEL STANDARD; PRT; 1159 AA.  
AC Q09531; 1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE HYPOTHETICAL 127.4 KDA PROTEIN F07F6.4 IN CHROMOSOME III.  
GN F07F6.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Chisoe S.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE GCS1/GLO3/SPS18 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U23486; AAC46777.1; -.  
DR WormPep; F07F6.4; CE01896.  
DR InterPro; IPR001164; -.  
DR Pfam; PF01412; ArfGap; 1.  
DR PRINTS; PR00405; REVINTRACTNG.  
KW Hypothetical protein; Zinc-finger; Nuclear protein; DNA-binding.  
FT ZN\_FING 28 51 C4-TYPE.  
FT DOMAIN 375 378 POLY-SER.  
FT DOMAIN 404 411 POLY-SER.  
FT DOMAIN 878 881 POLY-ASP.  
SQ SEQUENCE 1159 AA; 127429 MW; B0A7E5AABF815C63 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 1159;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1188 LKNGSSS 1194  
Db 1120 LKNGSSS 1126

RESULT 113  
APAF\_HUMAN STANDARD; PRT; 1194 AA.  
ID APAF\_HUMAN  
AC 014727; 1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE APOPTOTIC PROTEASE ACTIVATING FACTOR 1 (APAF-1).  
GN APAF1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RX MEDLINE=97410306; PubMed=9267021;  
RA Zou H., Henzel H.J., Liu X., Lutsch A., Wang X.;  
RT "Apaf-1, a human protein homologous to C. elegans CED-4, participates in cytochrome c-dependent activation of caspase-3.";  
RL Cell 90:405-413(1997).  
RN [2]

RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS) OF 1-97.  
RX MEDLINE=20013059; PubMed=10543941;

RA Vaughn D.E., Rodriguez J., Lazebnik Y., Joshua-Tor L.;  
 RT "Crystal structure of Apaf-1 caspase recruitment domain: an alpha-  
 RL helical Greek key fold for apoptotic signaling.";  
 CC J. Mol. Biol. 293:439-447(1999).  
 CC -!- FUNCTION: PARTICIPATES WITH CASPASE-9 (APAF-3) IN THE CYTOCHROME  
 CC C-DEPENDENT ACTIVATION OF CASPASE-3, LEADING TO APOPTOSIS. THIS  
 CC ACTIVATION REQUIRES ATP.  
 CC -!- SUBUNIT: CASPASE-9 AND APAF-1 BIND TO EACH OTHER VIA THEIR  
 CC RESPECTIVE NHE2-TERMINAL CED-3 HOMOLOGOUS DOMAINS IN THE PRESENCE  
 CC OF CYTOCHROME C AND ATP.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- TISSUE SPECIFICITY: UBIQUITOUS. HIGHEST LEVELS OF EXPRESSION IN  
 CC ADULT SPLEEN AND PERIPHERAL BLOOD LEUKOCYTES, AND IN FETAL BRAIN,  
 CC KIDNEY AND LUNG.  
 CC -!- SIMILARITY: CONTAINS 12 WD REPEATS (TRP-ASP DOMAINS).  
 CC -!- SIMILARITY: CONTAINS 1 CARD DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF013263; AAC51678.1; -;  
 DR PDB; 1CV5; 01-DEC-99.  
 DR MIM; 602233; -;  
 DR InterPro; IPR001315; -;  
 DR InterPro; IPR001680; -;  
 DR InterPro; IPR002182; -;  
 DR Pfam; PF00619; CARD; 1.  
 DR Pfam; PF00931; NB-ARC; 1.  
 DR Pfam; PF00400; WD40; 10.  
 DR PRINTS; PRO0320; GPROTEINBRPT.  
 DR PROSITE; PS00209; CARD; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE; PS00082; WD\_REPEATS\_2; 9.  
 DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
 KW Apoptosis; Repeat; WD repeat; 3D-structure.  
 FT DOMAIN 1 89  
 FT REPEAT 602 641 WD 1.  
 FT REPEAT 644 683 WD 2.  
 FT REPEAT 686 727 WD 3.  
 FT REPEAT 730 769 WD 4.  
 FT REPEAT 785 823 WD 5.  
 FT REPEAT 826 865 WD 6.  
 FT REPEAT 905 944 WD 7.  
 FT REPEAT 947 986 WD 8.  
 FT REPEAT 988 1026 WD 9.  
 FT REPEAT 1029 1068 WD 10.  
 FT REPEAT 1071 1110 WD 11.  
 FT REPEAT 1121 1158 WD 12.  
 FT DOMAIN 95 98 POLY-SER.  
 SQ SEQUENCE 1194 AA; 135980 MW; A675EA102DDAAFB7 CRC64;  
 -----  
 Query Match 0.5%; Score 7; DB 1; Length 1194;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 935 DATSANE 941  
 DB 760 DATSANE 766  
 -----  
 RESULT 114  
 MGPC MYCPN  
 ID MGPC\_MYCPN STANDARD; PRT; 1218 AA.  
 AC Q50341;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE MGPC PROTEIN PRECURSOR.  
 GN MGPC OR MPN142 OR MP012.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=99211947; PubMed=2468577;  
 RA Inamine J.M., Loechele S., Hu P.C.;  
 RT "Analysis of the nucleotide sequence of the P1 operon of Mycoplasma  
 RT pneumoniae.";  
 RL Gene 73:175-183(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M21519; AAA88326.1; -;  
 DR EMBL; AE000002; AAB95660.1; -;  
 KW Cell adhesion; Signal; Membrane.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 1218 MGPC PROTEIN.  
 SQ SEQUENCE 1218 AA; 130456 MW; 4DA29BCE41538311 CRC64;  
 -----  
 Query Match 0.5%; Score 7; DB 1; Length 1218;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 967 QTLSLN 973  
 DB 336 QTLSLN 342  
 -----  
 RESULT 115  
 XDH\_RAT  
 ID XDH\_RAT STANDARD; PRT; 1330 AA.  
 AC P22985; Q63157;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE XANTHINE DEHYDROGENASE/OXIDASE [INCLUDES: XANTHINE DEHYDROGENASE  
 DE (EC 1.1.1.204) (XD); XANTHINE OXIDASE (EC 1.1.3.22) (XO) (XANTHINE  
 DE OXIDOREDUCTASE)].  
 GN XDH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=101116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Liver;  
 RX MEDLINE=90354396; PubMed=2387845;  
 RA Amaya Y., Yamazaki K.-I., Sato M., Noda K., Nishino T., Nishino T.;  
 RT "Proteolytic conversion of xanthine dehydrogenase from the  
 RT NAD-dependent type to the O2-dependent type. Amino acid sequence of  
 RT rat liver xanthine dehydrogenase and identification of the cleavage  
 RT sites of the enzyme protein during irreversible conversion by

RT tryptsin.";  
 RL J. Biol. Chem. 265:14170-14175(1990).  
 RN [2]  
 RP SEQUENCE OF 1-54 FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RX MEDLINE=94268906; PubMed=8208609;  
 RA Chow C.W., Clark M., Rinaldo J., Chalkley R.;  
 RT "Identification of the rat xanthine dehydrogenase/oxidase promoter.";  
 RL Nucleic Acids Res. 22:1846-1854(1994).  
 CC -1- FUNCTION: THIS ENZYME CAN BE CONVERTED FROM THE DEHYDROGENASE FORM  
 CC (D) TO THE OXIDASE FORM (O) IRREVERSIBLY BY PROTEOLYSIS OR  
 CC REVERSIBLY THROUGH THE OXIDATION OF SULFHYDRYL GROUPS.  
 CC -1- CATALYTIC ACTIVITY: XANTHINE + NAD(+) + H(2)O - URATE + NADH.  
 CC -1- CATALYTIC ACTIVITY: XANTHINE + H(2)O + O(2) - URATE + H(2)O(2).  
 CC -1- COFACTOR: FAD, MOLYBDOPTERIN, AND TWO 2FE-2S CLUSTERS.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: PEROXISOMAL.  
 CC -1- INDUCTION: BY INTERFERON.  
 CC -1- SIMILARITY: TO OTHER XANTHINE DEHYDROGENASES, AND LIMITED TO  
 CC OTHER EUKARYOTIC MOLYBDOPTERIN ENZYMES SUCH AS NITRATE REDUCTASE  
 CC AND SULFITE OXIDASE.  
 CC -1- SIMILARITY: TO 2FE-2S FERREDOXINS IN THE N-TERMINAL DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J05579; AAA42349.1; -;  
 DR EMBL; U08122; AAA18869.1; -;  
 DR EMBL; U08120; AAA18869.1; JOINED.  
 DR EMBL; U08121; AAA18869.1; JOINED.  
 DR PIR; A37810; A37810.  
 DR HSP; Q46509; IALO.  
 DR InterPro: IPR000564; -;  
 DR InterPro: IPR000572; -;  
 DR InterPro: IPR000674; -;  
 DR InterPro: IPR001041; -;  
 DR InterPro: IPR002346; -;  
 DR InterPro: IPR002888; -;  
 DR Pfam; PF01315; Ald\_Xan\_dh\_c; 1.  
 DR Pfam; PF00941; dehydrog\_molyb; 1.  
 DR Pfam; PF00111; fer2; 1.  
 DR Pfam; PF01799; fer2\_2; 1.  
 DR PROSITE; PS00197; 2FE2S\_FERREDOXIN; 1.  
 DR PROSITE; PS00559; MOLYBDOPTERIN\_EUK; 1.  
 KW Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Iron-sulfur.  
 FT INIT\_MET 0 -;  
 FT METAL 36 36 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 42 42 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 47 47 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 50 50 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 SQ SEQUENCE 1330 AA; 146111 MW; A3DD206B9D74E565 CRC64;  
 Query Match 0.5%; Score 7; DB 1; Length 1330;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 711 SATGFYK 717  
 |||||  
 Db 1127 SATGFYK 1133  
 RESULT 116  
 XDH\_MOUSE  
 ID XDH\_MOUSE STANDARD; PRT; 1335 AA.  
 AC Q00519;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE XANTHINE DEHYDROGENASE/OXIDASE [INCLUDES: XANTHINE DEHYDROGENASE  
 DE (EC 1.1.1.204) (XD); XANTHINE OXIDASE (EC 1.1.3.22) (XO) (XANTHINE  
 DE OXIDOREDUCTASE)].  
 GN XDH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-129/SV; TISSUE=Spleen;  
 RX MEDLINE=95137585; PubMed=7835888;  
 RA Cazzaniga G., Terao M., Lo Schiavo P., Galbiati F., Segalla F.,  
 RA Seldin M.F., Garattini E.;  
 RT "Chromosomal mapping, isolation, and characterization of the mouse  
 RL xanthine dehydrogenase gene.";  
 RL Genomics 23:390-402(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL; TISSUE=Liver;  
 RX MEDLINE=92272690; PubMed=1590774;  
 RA Terao M., Cazzaniga G., Ghezzi P., Bianchi M., Falciani F.,  
 RA Perani P., Garattini E.;  
 RT "Molecular cloning of a cDNA coding for mouse liver xanthine  
 RT dehydrogenase. Regulation of its transcript by interferons in vivo.";  
 RL Biochem. J. 283:863-870(1992).  
 CC -1- FUNCTION: THIS ENZYME CAN BE CONVERTED FROM THE DEHYDROGENASE FORM  
 CC (D) TO THE OXIDASE FORM (O) IRREVERSIBLY BY PROTEOLYSIS OR  
 CC REVERSIBLY THROUGH THE OXIDATION OF SULFHYDRYL GROUPS.  
 CC -1- CATALYTIC ACTIVITY: XANTHINE + NAD(+) + H(2)O - URATE + NADH.  
 CC -1- CATALYTIC ACTIVITY: XANTHINE + H(2)O + O(2) - URATE + H(2)O(2).  
 CC -1- COFACTOR: FAD, MOLYBDOPTERIN, AND TWO 2FE-2S CLUSTERS.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: PEROXISOMAL.  
 CC -1- INDUCTION: BY INTERFERON.  
 CC -1- SIMILARITY: TO OTHER XANTHINE DEHYDROGENASES, AND LIMITED TO  
 CC OTHER EUKARYOTIC MOLYBDOPTERIN ENZYMES SUCH AS NITRATE REDUCTASE  
 CC AND SULFITE OXIDASE.  
 CC -1- SIMILARITY: TO 2FE-2S FERREDOXINS IN THE N-TERMINAL DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X75129; CAA52997.1; -;  
 DR EMBL; X75128; CAA52997.1; JOINED.  
 DR EMBL; X75127; CAA52997.1; JOINED.  
 DR EMBL; X75126; CAA52997.1; JOINED.  
 DR EMBL; X75125; CAA52997.1; JOINED.  
 DR EMBL; X75124; CAA52997.1; JOINED.  
 DR EMBL; X75123; CAA52997.1; JOINED.  
 DR EMBL; X75122; CAA52997.1; JOINED.  
 DR EMBL; X75121; CAA52997.1; JOINED.  
 DR EMBL; X75120; CAA52997.1; JOINED.  
 DR EMBL; X75119; CAA52997.1; JOINED.  
 DR EMBL; X75130; CAA52997.1; JOINED.  
 DR EMBL; X75131; CAA52997.1; JOINED.  
 DR EMBL; X75132; CAA52997.1; JOINED.  
 DR EMBL; X75133; CAA52997.1; JOINED.  
 DR EMBL; X75134; CAA52997.1; JOINED.  
 DR EMBL; X75135; CAA52997.1; JOINED.  
 DR EMBL; X75136; CAA52997.1; JOINED.  
 DR EMBL; X75137; CAA52997.1; JOINED.  
 DR EMBL; X75138; CAA52997.1; JOINED.  
 DR EMBL; X75139; CAA52997.1; JOINED.  
 DR EMBL; X75140; CAA52997.1; JOINED.  
 DR EMBL; X75141; CAA52997.1; JOINED.  
 DR EMBL; X75142; CAA52997.1; JOINED.

CC EMBL; X75143; CAA52997.1; JOINED.  
 CC EMBL; X75151; CAA52997.1; JOINED.  
 CC EMBL; X75152; CAA52997.1; JOINED.  
 CC EMBL; X75153; CAA52997.1; JOINED.  
 CC EMBL; X75154; CAA52997.1; JOINED.  
 CC EMBL; X75144; CAA52997.1; JOINED.  
 CC EMBL; X75145; CAA52997.1; JOINED.  
 CC EMBL; X75146; CAA52997.1; JOINED.  
 CC EMBL; X75147; CAA52997.1; JOINED.  
 CC EMBL; X75148; CAA52997.1; JOINED.  
 CC EMBL; X75149; CAA52997.1; JOINED.  
 CC EMBL; X75150; CAA52997.1; JOINED.  
 CC EMBL; X62932; CAA52997.1; JOINED.  
 CC PIR; S22419; S22419.  
 CC HSSP; Q46509; IALO.  
 CC MGB; MGI:98973; xdh.  
 CC InterPro; IPR000564; -.  
 CC InterPro; IPR000572; -.  
 CC InterPro; IPR000674; -.  
 CC InterPro; IPR001041; -.  
 CC InterPro; IPR002346; -.  
 CC InterPro; IPR002888; -.  
 CC Pfam; PF01315; Ald\_Xan\_dh\_C; 1.  
 CC Pfam; PF00941; dehydrog\_molyb; 1.  
 CC Pfam; PF00111; fer2; 1.  
 CC Pfam; PF01799; fer2; 2; 1.  
 CC PROSITE; PS00197; 2FE2S\_FERROXIN; 1.  
 CC PROSITE; PS00559; MOLYBDOPTERIN\_EUK; 1.  
 KW Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Iron-sulfur.  
 FT METAL 40 40 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 46 46 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 51 51 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 54 54 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 241 241 V -> I (IN REF. 2).  
 FT CONFLICT 621 621 T -> M (IN REF. 2).  
 SQ SEQUENCE 1335 AA; 146517 MW; 99C66FD8B42FB5E5 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 1335;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 711 SATGPK 717  
 Db 1131 SATGPK 1137  
 |||||

RESULT 117  
 HTK7\_HYDAT  
 ID HTK7\_HYDAT STANDARD; PRT; 1477 AA.  
 AC Q25197;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE PUTATIVE INSULIN-LIKE PEPTIDE RECEPTOR PRECURSOR (EC 2.7.1.112).  
 GN HTK7.  
 OS Hydra attenuata (Hydra) (Hydra vulgaris).  
 CC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;  
 CC Hydrozoa; Hydra.  
 CC NCBI\_taxid=6087;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Steele R.E., Mai N.H., Lieu P., Shenk M.A.;  
 RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: THIS RECEPTOR PROBABLY BINDS AN INSULIN RELATED PROTEIN  
 CC AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN DIVIDING EPITHELIAL CELLS.  
 CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

CC EMBL; M64612; AAA68205.1; -.  
 CC HSSP; P06213; IIRK.  
 CC InterPro; IPR000494; -.  
 CC InterPro; IPR000719; -.  
 CC InterPro; IPR001245; -.  
 CC InterPro; IPR001777; -.  
 CC InterPro; IPR002011; -.  
 CC InterPro; IPR002174; -.  
 CC Pfam; PF00757; Furin-like; 1.  
 CC Pfam; PF01030; Recep\_L\_domain; 1.  
 CC Pfam; PF00041; fn3; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Glycoprotein; ATP-binding; Phosphorylation; Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 1477 PUTATIVE INSULIN-LIKE PEPTIDE RECEPTOR.  
 FT DOMAIN 25 980 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 981 1001 POTENTIAL.  
 FT DOMAIN 1002 1477 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1044 ? PROTEIN KINASE.  
 FT NP\_BIND 1050 1058 ATP (BY SIMILARITY).  
 FT BINDING 1077 1077 ATP (BY SIMILARITY).  
 FT ACT\_SITE 1175 1175 BY SIMILARITY.  
 FT MOD\_RES 1201 1201 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 55 55 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 325 325 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 491 491 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 549 549 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 644 644 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 732 732 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 791 791 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 874 874 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 895 895 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 957 957 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 1477 AA; 168276 MW; 74ACDBA7C6DE1D41 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 1477;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 SKVTGEV 285  
 Db 493 SKVTGEV 499  
 |||||

RESULT 118  
 SYJ1\_HUMAN  
 ID SYJ1\_HUMAN STANDARD; PRT; 1575 AA.  
 AC Q43426; O43425;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE SYNAPTOJANIN 1 (EC 3.1.3.56) (SYNAPTIC INOSITOL-1,4,5-TRISPHOSPHATE 5-  
 DE PHOSPHATASE 1).  
 GN SYN1.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=98088905; PubMed=9428629;  
 RA Haffner C., Takei K., Chen H., Ringstad N., Hudson A., Butler M.H.,  
 RA Salcini A.E., Di Fiore P.P., De Camilli P.;  
 RT "Synaptotagmin 1: localization on coated endocytic intermediates in  
 RT nerve terminals and interaction of its 170 kDa isoform with Eps15.";  
 RL FEBS Lett. 419:175-180(1997).  
 CC -1- FUNCTION: INOSITOL 5-PHOSPHATASE WHICH HAS A ROLE IN CLATHRIN-  
 CC MEDIATED ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: D-MYO-INOSITOL 1,4,5-TRISPHOSPHATE + H(2)O =  
 CC D-MYO-INOSITOL 1,4-BISPHOSPHATE + PHOSPHATE.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG ISOFORM/SYNAPTOTAGMIN-  
 CC 170 (SHOWN HERE) AND A SHORT ISOFORM/SYNAPTOTAGMIN-145; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: CONCENTRATED AT CLATHRIN-COATED ENDOCYTIC  
 CC INTERMEDIATES IN NERVE TERMINALS. THE LONG ISOFORM IS MORE  
 CC ENRICHED THAN THE SHORT ISOFORM IN DEVELOPING BRAIN AS WELL AS  
 CC NON-NEURONAL CELLS. THE SHORT ISOFORM IS VERY ABUNDANT IN NERVE  
 CC TERMINALS.  
 CC -1- DOMAIN: BINDS TO EPS15 (A CLATHRIN COAT-ASSOCIATED PROTEIN) VIA A  
 CC C-TERMINAL DOMAIN CONTAINING THREE ASN-PRO-PHE (NPF) REPEATS.  
 CC -1- DOMAIN: THE C-TERMINAL PROLINE-RICH REGION MEDIATES BINDING TO A  
 CC VARIETY OF SH3 DOMAIN-CONTAINING PROTEINS INCLUDING AMPHIPHYSIN,  
 CC SH3P4 AND GRB2.  
 CC -1- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE INOSITOL-1,4,5-  
 CC TRISPHOSPHATE 5-PHOSPHATASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 SAC1 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; AF009040; AAC51922.1; -;  
 DR EMBL; AF009039; AAC51921.1; -;  
 DR MIM; 604297; -;  
 DR InterPro: IPR000300; -;  
 DR Pfam; PF00783; IPPE; 1.  
 DR PROSITE; PS50102; RRM; 1.  
 DR Hydrolase; Alternative splicing; Repeat; Endocytosis; RNA-binding;  
 KW Multigene family.  
 FT DOMAIN 1 499  
 FT DOMAIN 500 899  
 FT DOMAIN 902 971  
 FT DOMAIN 900 1575  
 FT DOMAIN 1033 1036  
 FT DOMAIN 1108 1113  
 FT DOMAIN 1126 1129  
 FT DOMAIN 1487 1490  
 FT DOMAIN 1540 1546  
 FT DOMAIN 1396 1419  
 FT REPEAT 1396 1398  
 FT REPEAT 1406 1408  
 FT REPEAT 1417 1419  
 FT REPEAT 1417 1419  
 FT VARSPLIC 1306 1311  
 FT VARSPLIC 1312 1575  
 SEQUENCE 1575 AA; 173345 MW; 50646f6cc043b9e7 CRC64;  
 Query Match 0.5%; Score 7; DB 1; Length 1575;  
 Best Local Similarity 100.0%; Pred. No. 3e+02; 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;  
 QY 560 VKTNGIS 566

Db 1306 VKTNGIS 1312  
 RESULT 119  
 CO3\_EPTBU  
 ID CO3\_EPTBU STANDARD; PRT; 1620 AA.  
 AC P98094;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE COMPLEMENT C3 [CONTAINS: C3A ANAPHYLATOXIN] (FRAGMENT).  
 GN C3.  
 OS Eptaretus burgeri (Inshore hagfish).  
 OC Eukaryota; Chordata; Craniata; Hyperotreti; Myxiniiformes;  
 OC Myxiniidae; Eptaretinae; Eptaretus.  
 OX NCBI\_TaxID=7764;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=92192016; PubMed=1372251;  
 RA Ishiguro H., Kobayashi K., Suzuki M., Titani K., Tomonaga S.,  
 RA Kurosawa Y.;  
 RT "Isolation of a hagfish gene that encodes a complement component.";  
 RL EMBO J. 11:829-837(1992).  
 CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
 CC COMPLEMENT SYSTEM. AFTER ACTIVATION (C3B), IT CAN BIND COVALENTLY,  
 CC VIA ITS REACTIVE THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR  
 CC IMMUNE AGGREGATES. CYCLOSTOMATES C3 APPEARS TO REPRESENT THE  
 CC COMMON ANCESTOR OF MAMMALIAN C3 AND C4, SHOWING SIMILARITIES TO  
 CC BOTH PROTEINS.  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: CONTAINS 1 ANAPHYLATOXIN-LIKE DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; Z11595; CAA77677.1; -;  
 DR EMBL; Z11596; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S21045; S21045.  
 DR HSP; P01032; IC5A.  
 DR InterPro; IPR000020; -;  
 DR InterPro; IPR001134; -;  
 DR InterPro; IPR001599; -;  
 DR InterPro; IPR002890; -;  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M.N; 1.  
 DR Pfam; PF01821; ANATO; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR PROSITE; PS00477; ALPHA\_2-MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway; Plasma; Inflammatory response; Glycoprotein.  
 FT NON\_TER 1  
 FT CHAIN 1 1620  
 FT CHAIN <1 633  
 FT CHAIN 634 1336  
 FT CHAIN 1343 1620  
 FT PEPTIDE 634 714  
 FT DOMAIN 658 694  
 FT DOMAIN 1406 1416  
 FT DISULFID 526 780  
 FT DISULFID 593 628  
 FT DISULFID 658 686  
 FT DISULFID 659 693  
 FT DISULFID 672 694  
 FT DISULFID 836 1474  
 FT DISULFID 1062 1114  
 FT DISULFID 1321 1451  
 COMPLEMENT C3.  
 BETA CHAIN (BY SIMILARITY).  
 ALPHA CHAIN (BY SIMILARITY).  
 GAMMA CHAIN (BY SIMILARITY).  
 C3A ANAPHYLATOXIN (BY SIMILARITY).  
 ANAPHYLATOXIN-LIKE.  
 PROPERDIN-BINDING.  
 INTERCHAIN (BY SIMILARITY).  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.

Query Match 0.5%; Score 7; DB 1; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 962 KTSGLQT 968  
DB 1170 KTSGLQT 1176

RESULT 120  
MLP1\_YEAST STANDARD; PRT; 1875 AA.

AC Q02455;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE MYOSIN-LIKE PROTEIN MLP1.  
GN MLP1 OR YKR095W OR YKR415.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=93247549; PubMed=8483450;  
RA Koeilling R., Nguyen T., Chen E.Y., Botstein D.;  
RT "A new yeast gene with a myosin-like heptad repeat structure.";  
RL Mol. Gen. Genet. 237:359-369(1993).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94205265; PubMed=8154186;  
RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,  
RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;  
RT "The complete sequence of a 15,820 bp segment of Saccharomyces  
cerevisiae chromosome XI contains the UBI2 and MLP1 genes and three  
new open reading frames.";  
RL Yeast 9:1349-1354(1993).  
-!- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA  
REPAIR.  
-!- SIMILARITY: SOME, TO THE TPR ONCOGENE.  
-!- CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MLP1".

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: L01992; AAA34783.1;  
EMBL: X73541; CAAS1948.1;  
EMBL: Z28320; CAAB2174.1;  
PIR: S38173; S38173.  
SGD: S0001803; MLP1.  
Myosin; Heptad repeat pattern: Coiled coil; DNA repair.  
DOMAIN 69 487 COILED COIL (POTENTIAL).  
DOMAIN 531 1678 COILED COIL (POTENTIAL).  
DOMAIN 1834 1866 COILED COIL (POTENTIAL).  
CONFLICT 301 301 R -> A (IN REF. 1).  
SEQUENCE 1875 AA; 218455 MW; 683A0D34C9066867 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 1875;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1260 ELKLAKE 1266  
DB 1271 ELKLAKE 1277

RESULT 121  
FAS2\_CANAL STANDARD; PRT; 1885 AA.

AC P43098;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE FATTY ACID SYNTHASE, SUBUNIT ALPHA (EC 2.3.1.86) [INCLUDES:  
EC 1.1.1.100; EC 2.3.1.41].  
GN FAS2.  
OS Candida albicans (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=5476;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=4918;  
RX MEDLINE=95255657; PubMed=7737507;  
RA Southard S.B., Cihlar R.L.;  
RT "Analysis and expression of the Candida albicans FAS2 gene.";  
RL Gene 156:133-138(1995).  
-!- FUNCTION: FATTY ACID SYNTHASE CATALYZES THE FORMATION OF  
LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.  
THE ALPHA SUBUNIT CONTAINS DOMAINS FOR: ACYL CARRIER PROTEIN.  
3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE, AND 3-OXOACYL-[ACYL-  
CARRIER-PROTEIN] SYNTHASE.  
-!- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH -  
NADP(+) = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.  
-!- SUBUNIT: FATTY ACID SYNTHETASE ARE [ALPHA(6)BETA(6)] HEXAMERS OF  
TWO MULTIFUNCTIONAL SUBUNITS (ALPHA & BETA).  
-!- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + MALONYL-[ACYL-  
CARRIER PROTEIN] = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + CO(2) +  
[ACYL-CARRIER PROTEIN].  
-!- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +  
NADP(+) = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.  
-!- SUBUNIT: FATTY ACID SYNTHETASE ARE [ALPHA(6)BETA(6)] HEXAMERS OF  
TWO MULTIFUNCTIONAL SUBUNITS (ALPHA & BETA).  
-!- SIMILARITY: TO THE FATTY ACID SYNTHETASE, SUBUNIT ALPHA FROM  
OTHER FUNGI.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: L29063; AAA34345.1;  
InterPro: IPR000255;  
InterPro: IPR000794;  
InterPro: IPR002582;  
Pfam: PF01648; ACPS; 1.  
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.  
DR PROSITE: PS00606; B\_KETOACYL SYNTHASE; 1.  
Fatty acid biosynthesis; Multifunctional enzyme; Oxidoreductase;  
KW Transferase: NADP; Phosphopantetheine.  
FT DOMAIN 1 ? ACYL CARRIER.  
FT DOMAIN ? ? BETA-KETOACYL REDUCTASE.  
FT DOMAIN ? 1885 BETA-KETOACYL SYNTHASE.  
FT BINDING 181 181 PHOSPHOPANTETHEINE (BY SIMILARITY).  
FT ACT\_SITE 1304 1304 BETA-KETOACYL SYNTHASE (BY SIMILARITY).  
SQ SEQUENCE 1885 AA; 207588 MW; 4835D57F362372E0 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 1885;



Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 346 NDKNESA 352  
| | | | |  
Db 1553 NDKNESA 1559

# RESULT 122

FAS2\_YEAST  
ID FAS2\_YEAST STANDARD; PRT; 1894 AA.  
AC P19097;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE FATTY ACID SYNTHASE, SUBUNIT ALPHA (EC 2.3.1.86) [INCLUDES:  
DE EC 1.1.1.100; EC 2.3.1.41].  
GN FAS2 OR YPL231W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-88315020; PubMed-2900835;  
RA Mohamed A.H., Chirala S.S., Mody N.H., Huang W.Y., Wakil S.J.;  
RT "Primary structure of the multifunctional alpha subunit protein of  
RT yeast fatty acid synthase derived from FAS2 gene sequence.";  
RL J. Biol. Chem. 263:12315-12325(1988).  
RN [2]  
RP MUTAGENESIS OF GLY-1257.  
RC STRAIN-S288C;  
RX MEDLINE-94316198; PubMed-8041367;  
RA Inokoshi J., Tomoda H., Hashimoto H., Watanabe A., Takeshima H.,  
RA Omura S.;  
RT "Cerulein-resistant mutants of Saccharomyces cerevisiae with an  
RT altered fatty acid synthase gene.";  
RL Mol. Gen. Genet. 244:90-96(1994).

CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF  
CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.  
CC THE ALPHA SUBUNIT CONTAINS DOMAINS FOR: ACYL CARRIER PROTEIN,  
CC 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE, AND 3-OXOACYL-[ACYL-  
CC CARRIER-PROTEIN] SYNTHASE. THIS SUBUNIT COORDINATES THE BINDING  
CC OF THE SIX BETA SUBUNITS TO THE ENZYME COMPLEX.  
CC -1- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH =  
CC LONG-CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2N NADP(+).  
CC -1- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + MALONYL-[ACYL-  
CC CARRIER PROTEIN] = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + CO(2) +  
CC [ACYL-CARRIER PROTEIN].  
CC -1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +  
CC NADP(+) = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.  
CC -1- SUBUNIT: FATTY ACID SYNTHETASE ARE [ALPHA(6)BETA(6)] HEXAMERS OF  
CC TWO MULTIFUNCTIONAL SUBUNITS (ALPHA & BETA).  
CC -1- SIMILARITY: TO THE FATTY ACID SYNTHETASE, SUBUNIT ALPHA FROM  
CC OTHER FUNGI.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; J03936; AAA34601.1;  
CC PIR; A31107; A31107.  
CC SGD; S0006152; FAS2.  
CC InterPro; IPR000255;  
CC InterPro; IPR000794;  
CC InterPro; IPR002582;  
CC Pfam; PF01648; ACPS; 1.  
CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.  
CC PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 1.

KW Fatty acid biosynthesis; Multifunctional enzyme; Oxidoreductase;  
KW Transferase; NADP; Phosphopantetheine.

FT DOMAIN 1 ? ACYL CARRIER.  
FT TRANSFERASE 682 881 BETA-KETOACYL REDUCTASE.  
FT DOMAIN 180 180 BETA-KETOACYL SYNTHASE.  
FT BINDING 180 180 PHOSPHOPANTETHEINE (BY SIMILARITY).  
FT ACT\_SITE 1312 1312 BETA-KETOACYL SYNTHASE (BY SIMILARITY).  
FT MUTAGEN 1257 1257 G->S: CERULENIN-RESISTANCE.  
SQ SEQUENCE 1894 AA; 208098 MW; AEC0B2269DEE833 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 1894;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 346 NDKNESA 352  
| | | | |  
Db 1556 NDKNESA 1562

# RESULT 123

TUD\_DROME  
ID TUD\_DROME STANDARD; PRT; 2515 AA.  
AC P25823;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE MATERNAL TUDOR PROTEIN.  
GN TUD.

OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92038995; PubMed-1936993;  
RA Golumbeski G.S., Bardsley A., Tax F., Boswell R.E.;  
RT "Tudor, a posterior-group gene of Drosophila melanogaster, encodes a  
RT novel protein and an mRNA localized during mid-oogenesis.";  
RL Genes Dev. 5:2060-2070(1991).  
CC -1- FUNCTION: REQUIRED DURING OOCYTES FOR THE FORMATION OF  
CC PRIMORDIAL GERM CELLS AND FOR NORMAL ABDOMINAL SEGMENTATION.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT THE LIFE CYCLE.  
CC -1- MISCELLANEOUS: THE TUD MRNA ACCUMULATES WITHIN THE POSTERIOR  
CC REGION OF THE DEVELOPING OOCYTE DURING THE EARLY TO MIDDLE STAGES  
CC OF OOCYTES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X62420; CAA44286.1;  
CC PIR; A41519; A41519.  
CC FlyBase; FBgn0003891; tud.  
CC InterPro; IPR001097;  
CC Pfam; PF00567; TUDOR; 10.  
CC Developmental protein.  
SQ SEQUENCE 2515 AA; 285236 MW; 683C100AD308BADA CRC64;

Query Match 0.5%; Score 7; DB 1; Length 2515;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1004 DQKFASL 1010  
| | | | |  
Db 180 DQKFASL 186



```

RESULT 124
POLG_YEFV1 STANDARD; PRT; 3411 AA.
AC P03314; 042028;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX
DE PROTEIN (ENVELOPE PROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL
DE PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED
DE RNA POLYMERASE (EC 2.7.7.48) (NS5)].
OS Yellow fever virus (strain 17D).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=11090;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=85272570; PubMed=4023707;
RA Rice C.M., Lencches E.M., Eddy S.R., Shin S.J., Sheets R.L.,
RA Strauss J.H.;
RT "Nucleotide sequence of yellow fever virus: implications for
RT flavivirus gene expression and evolution.";
RL Science 229:726-733(1985).
CC -!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X03700; CAA27332.1; -
DR PIR; A03914; GNWVY.
DR HSP; P14336; 1SVB.
DR MEROPS; S07.001; -.
DR InterPro; IPR0000069; -.
DR InterPro; IPR000208; -.
DR InterPro; IPR000336; -.
DR InterPro; IPR000404; -.
DR InterPro; IPR000487; -.
DR InterPro; IPR000752; -.
DR InterPro; IPR001122; -.
DR InterPro; IPR001157; -.
DR InterPro; IPR001528; -.
DR InterPro; IPR001850; -.
DR InterPro; IPR002535; -.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF00948; Flavi_NS1; 1.
DR Pfam; PF01005; Flavi_NS2A; 1.
DR Pfam; PF01002; Flavi_NS2B; 1.
DR Pfam; PF01350; Flavi_NS4A; 1.
DR Pfam; PF01349; Flavi_NS4B; 1.
DR Pfam; PF00972; Flavi_NS5; 1.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF00949; Flavi_helicase; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
KW Transmembrane; Nonstructural protein.
FT INIT_MET 1 1
FT CELLULAR AMINOPEPTIDASE.
FT CHAIN 1 121
FT PROPEP 122 210

```

```

FT CHAIN 211 285
FT CHAIN 286 778
FT CHAIN 779 1130
FT CHAIN 1131 1354
FT CHAIN 1355 1484
FT CHAIN 1485 2107
FT CHAIN 2108 2256
FT CHAIN 2257 2506
FT CHAIN 2507 3411
FT CHAIN 3411 3411
FT TRANSMEM 249 269
FT TRANSMEM 271 285
FT TRANSMEM 740 753
FT TRANSMEM 755 778
FT TRANSMEM 1159 1180
FT DOMAIN 383 396
FT NP_BIND 1682 1689
FT SITE 1773 1776
FT DISULFID 288 315
FT DISULFID 345 401
FT DISULFID 359 390
FT DISULFID 377 406
FT DISULFID 467 568
FT DISULFID 585 615
FT CARBOHYD 134 134
FT CARBOHYD 150 150
FT CARBOHYD 908 908
FT CARBOHYD 986 986
FT CARBOHYD 2320 2320
FT CARBOHYD 2346 2346
FT CARBOHYD 2467 2467
FT SEQUENCE 3411 AA; 379512 MW; 680E0FACD23DCFA6 CRC64;

Query Match 0.5%; Score 7; DB 1; Length 3411;
Best Local Similarity 100.0%; Pred. No. 6,1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22
| | | | |
Db 1455 SLALVGA 1461

RESULT 125
POLG_YEFV2 STANDARD; PRT; 3411 AA.
AC P19901;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX
DE PROTEIN (ENVELOPE PROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL
DE PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED
DE RNA POLYMERASE (EC 2.7.7.48) (NS5)].
OS Yellow fever virus (strain Pasteur 17D-204).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=11091;
RN SEQUENCE FROM N.A.
RP MEDLINE=89282413; PubMed=2734112;
RA Dupuy A., Despres P., Cahour A., Girard M., Bouloy M.;
RT "Nucleotide sequence comparison of the genome of two 17D-204 yellow
RT fever vaccines.";
RL Nucleic Acids Res. 17:3989-3989(1989).
CC -!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X15062; -; NOT\_ANNOTATED\_CDS.

PIR: S07757; GNAVYP.

HSSP: P14336; LSVB.

MEROPS: S07.001; -.

InterPro: IPR000069; -.

InterPro: IPR000208; -.

InterPro: IPR000336; -.

InterPro: IPR000404; -.

InterPro: IPR000487; -.

InterPro: IPR000752; -.

InterPro: IPR001122; -.

InterPro: IPR001157; -.

InterPro: IPR001528; -.

InterPro: IPR001850; -.

InterPro: IPR002535; -.

Pfam: PF01004; Flavi\_M; 1.

Pfam: PF00948; Flavi\_NS1; 1.

Pfam: PF01005; Flavi\_NS2A; 1.

Pfam: PF01002; Flavi\_NS2B; 1.

Pfam: PF01350; Flavi\_NS2A; 1.

Pfam: PF01349; Flavi\_NS4B; 1.

Pfam: PF00972; Flavi\_NS5; 1.

Pfam: PF01003; Flavi\_capsid; 1.

Pfam: PF00869; Flavi\_glycoprot; 1.

Pfam: PF00949; Flavi\_helicase; 1.

Pfam: PF01570; Flavi\_propep; 1.

Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;

Core protein; Coat protein; Helicase; ATP-binding;

Transmembrane; Nonstructural

INIT\_MET 1 1

REMOVED FROM CAPSID PROTEIN C BY THE

CELLULAR AMINOPEPTIDASE.

CAPSID PROTEIN C.

ENVELOPE GLYCOPROTEIN M.

MAJOR ENVELOPE PROTEIN E.

NONSTRUCTURAL PROTEIN NS1.

NONSTRUCTURAL PROTEIN NS2A.

NONSTRUCTURAL PROTEIN NS2B.

HELICASE (NS3).

NONSTRUCTURAL PROTEIN NS4A.

NONSTRUCTURAL PROTEIN NS4B.

RNA-DIRECTED RNA POLYMERASE (NS5).

ATP (POTENTIAL).

DEAH BOX.

POTENTIAL.

POTENTIAL.

POTENTIAL.

POTENTIAL.

POTENTIAL.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
|||||  
DB 1455 SLALVGA 1461

RESULT 126

T4C\_PARTE

ID T4C\_PARTE STANDARD; PRT; 27 AA.

AC Q27176;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE TRICHOYST MATRIX PROTEIN T4-C (SECRETORY GRANULE PROTEIN T4-C)

DE (TMP 4-C) (FRAGMENT).

GN T4C.

OS Paramecium tetraurelia.

OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;

OC Paramecium.

OX NCBI\_TaxID=5888;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=D4-2;

RX MEDLINE=96059477; PubMed=7579685;

RA Madeddu L., Gautier M.-C., Vayssie L., Houari A., Sperling L.;

RT "A large multigene family codes for the polypeptides of the

RT crystalline trichocyst matrix in Paramecium.";

RL Mol. Biol. Cell 6:649-659(1995).

RN [2]

RP PARTIAL SEQUENCE.

RC STRAIN=D4-2;

RX MEDLINE=95119139; PubMed=7819344;

RA Madeddu L., Gautier M.-C., le Caer J.P., de Loubresse N., Sperling L.;

RT "Protein processing and morphogenesis of secretory granules in

RT Paramecium.";

RL Biochimie 76:329-335(1994).

CC -I- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE

CC TRICHOYST MATRIX.

CC -I- SUBCELLULAR LOCATION: TRICHOYST. THESE ARE ARCHITECTURALLY

CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,

CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.

CC -I- SIMILARITY: BELONGS TO THE TMP FAMILY.

CC -I- DATABASE: NAME-Protein Spotlight;

CC NOTE-Issue 3 of October 2000;

CC WWW="http://www.expasy.ch/spotlight/articles/sptit003.html"

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U27514; AAA92614.1; -.

CC Polyprotein; Structural protein; Multigene family.

FT NON\_TER 1 1

FT NON\_TER 27 27

FT SEQUENCE 27 AA; 2837 MW; 731046E30185A542 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 27;

Best Local Similarity 100.0%; Pred. No. 86;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 LNNIAS 958

|||||

DB 10 LNNIAS 15

RESULT 127

Query Match 0.5%; Score 7; DB 1; Length 3411;

PT1\_ENTFA STANDARD; PRT; 29 AA.

AC P23530;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last annotation update)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE PHOSPHOENOLPYRUVATE-PROTEIN PHOSPHOTRANSFERASE (EC 2.7.3.9)

DE (PHOSPHOTRANSFERASE SYSTEM, ENZYME I) (FRAGMENT).

GN PTSI.

OS Enterococcus faecalis (Streptococcus faecalis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;

OC Enterococcus.

OX NCBI\_TaxID=1351;

RN [1]

RP SEQUENCE.

RX MEDLINE=85199851; PubMed=3922407;

RA Albert C.-A., Frank R., Stueber K., Deutscher J., Hengstenberg W.;

RT "Phosphoenolpyruvate-dependent protein kinase enzyme I of

RT Streptococcus faecalis: purification and properties of the enzyme and

RT characterization of its active center.";

RL Biochemistry 24:959-964(1985).

CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT

CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE

CC -TRANSPORT SYSTEM. ENZYME I TRANSFERS THE PHOSPHORYL GROUP FROM

CC PHOSPHOENOLPYRUVATE (PEP) TO THE PHOSPHORYL CARRIER PROTEIN (HPR).

CC ENZYME I IS COMMON TO ALL PTS.

CC -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + PROTEIN HISTIDINE =

CC PYRUVATE + PROTEIN N-PHOSPHOHISTIDINE.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -!- SIMILARITY: BELONGS TO THE PEP-UTILIZING ENZYMES FAMILY.

DR PIR; A22018; A22018.

DR HSP; P08839; 2E2B.

DR InterPro: IPR000121; .

DR Pfam: PF00391; PEP-utilizers; 1. PARTIAL.

DR PROSITE; PS00742; PEP\_ENZYMES\_2; PARTIAL.

DR PROSITE; PS00370; PEP\_ENZYMES\_PHOS\_SITE; 1.

KW Phosphotransferase system; Transferase; Kinase; Sugar transport;

KW Phosphorylation.

FT NON\_TER 1 1

FT MOD\_RES 12 12 PHOSPHORYLATION.

FT SEQUENCE 29 AA; 2999 MW; EGAC7E23E35BBAE8 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 29;

Best Local Similarity 100.0%; Pred. No. 92;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 IPAIVG 46

DB 22 IPAIVG 27

RESULT 128

RL16\_AQUYP

ID RL16\_AQUYP STANDARD; PRT; 49 AA.

AC Q92143;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE 50S RIBOSOMAL PROTEIN L16 (FRAGMENT).

GN RPLP OR RPL16.

OS Aquifex pyrophilus.

OC Bacteria; Aquificales; Aquificaceae; Aquifex.

OX NCBI\_TaxID=2714;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=KOL5A / DSM 6858;

RA Bocchetta M., Sanangelantoni A.M., Cammarano P.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS

CC LOCATED AT THE A SITE OF THE PEPTIDYLTRANSFERASE CENTER

CC (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AF040100; AAD08792.1; .

DR InterPro: IPR000114; .

DR Pfam: PF00252; Ribosomal\_L16; 1.

DR PRINTS; PR00060; RIBOSOMAL\_L16.

DR PROSITE; PS00586; RIBOSOMAL\_L16\_1; PARTIAL.

DR PROSITE; PS00701; RIBOSOMAL\_L16\_2; PARTIAL.

KW Ribosomal protein; rRNA-binding.

FT NON\_TER 49 49

FT SEQUENCE 49 AA; 5475 MW; 12EBDD60041A1342 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 49;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 633 NKLAFG 638

DB 17 NKLAFG 22

RESULT 129

DDR2\_YEAST

ID DDR2\_YEAST STANDARD; PRT; 61 AA.

AC P89113;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE DDR2 PROTEIN PRECURSOR.

GN DDR2 OR YOL052BC.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97127438; PubMed=8954934;

RA Kobayashi N., McClanahan T.K., Simon J.R., Treger J.M., McEntee K.;

RT "Structure and functional analysis of the multistress response gene

RT DDR2 from Saccharomyces cerevisiae.";

RL Biochem. Biophys. Res. Commun. 229:540-547(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / FY73;

RX MEDLINE=96381248; PubMed=8789261;

RA Mannhaupt G., Vetter I., Schwarzllose C., Mitzel S., Feldmann H.;

RT "Analysis of a 26 kb region on the left arm of yeast chromosome XV.";

RL Yeast 12:67-76(1996).

RN [3]

RP SEQUENCE FROM N.A.

RA Ansoorge W., Benes V., Rechmann S., Schwager C., Teodoru C.,

RA Voss H., Wiemann S.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE RESPONSE OF CELLS TO

CC DIVERSE ENVIRONMENTAL STRESSES.

CC -!- INDUCTION: BY MULTISTRESS. EXPRESSION IS CONTROLLED BY THE MSN2

CC AND MSN4 TRANSCRIPTIONAL REGULATORS.

CC -!- SIMILARITY: TO YEAST HOR7.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

```

CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U82215; AAB40031.1; -
DR      EMBL; 274794; CAA99059.1; -
DR      EMBL; 274795; CAA99061.1; -
DR      SGD; S0005413; DDR2.
KW      Signal.
FT      SIGNAL      1      22      POTENTIAL.
FT      CHAIN      23      61      DDR2 PROTEIN.
FT      CARBOHYD      24      24      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD      27      27      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE      61 AA; 5955 MW; F89BAAF9F466BD05 CRC64;

Query Match      0.5%; Score 6; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      851 NNTSNA 856
Db      27 NNTSNA 32

RESULT 130
RK35_PORPU
ID      RK35_PORPU      STANDARD;      PRT;      65 AA.
AC      P51270;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      CHLOROPLAST 50S RIBOSOMAL PROTEIN L35.
GN      RPL35.
OS      Porphyra purpurea.
OG      Chloroplast.
OC      Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX      NCBI_TaxID=2787;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=AVONPORT;
RA      Reith M.E., Munnholland J.;
RT      "Complete nucleotide sequence of the Porphyra purpurea chloroplast
      genome."
RL      Plant Mol. Biol. Rep. 13:333-335(1995).
CC      -1- SIMILARITY: BELONGS TO THE L35P FAMILY OF RIBOSOMAL PROTEINS.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U39804; AAC08156.1; -
DR      Mendel; 10332; PORPU; rpl35; 1.
DR      InterPro; IPR001706; -
DR      Pfam; PF01632; Ribosomal_L35p; 1.
DR      PRINTS; PR00064; RIBOSOMALL35.
DR      PROSITE; PS00936; RIBOSOMAL_L35; 1.
KW      Ribosomal protein; Chloroplast.
SQ      SEQUENCE      65 AA; 7371 MW; 365F7F52C0144C1B CRC64;

Query Match      0.5%; Score 6; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      192 AKNTSI 197
Db      55 AKNTSI 60

RESULT 131

```

```

VLYS_BPP21
ID      VLYS_BPP21      STANDARD;      PRT;      71 AA.
AC      P27360;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      01-JUN-1994 (Rel. 29, Last annotation update)
DE      LYSIS PROTEIN S.
GN      S.
OS      Bacteriophage P21 (Bacteriophage 21).
OC      Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC      Lambda phage group.
OX      NCBI_TaxID=10711;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=91210180; PubMed=2019562;
RA      Bonovich M.T., Young R.;
RT      "Dual start motif in two lambdoid S genes unrelated to lambda S.";
RL      J. Bacteriol. 173:2897-2905(1991).
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M65239; AAA32349.1; -
DR      PIR; S22905; S22905.
KW      Phage lysis protein.
SQ      SEQUENCE      71 AA; 7893 MW; 8690A8F25234A3E2 CRC64;

Query Match      0.5%; Score 6; DB 1; Length 71;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1049 YGTSAG 1054
Db      13 YGTSAG 18

RESULT 132
VLYS_ECOLI
ID      VLYS_ECOLI      STANDARD;      PRT;      71 AA.
AC      P77242;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-OCT-2000 (Rel. 40, Last annotation update)
DE      LYSIS PROTEIN S HOMOLOG FROM LAMBDOID PHOPHAGE DLP12.
GN      YBCR.
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA      Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA      Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL      Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: BELONGS TO THE LAMBDA PHAGE S PROTEIN FAMILY.
CC      -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE000161; AAC73655.1; -  
DR EMBL: U82598; AAB40750.1; -  
DR Ecogene: EGI3634; YBCK.  
KW Hypothetical protein; Phage lysis protein.  
SQ SEQUENCE 71 AA; 7778 MW; 9C013E2FE4361843 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 71;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 YGFSAG 1054  
Db 13 YGFSAG 18  
|||||

RESULT 133  
YF77\_HAEIN STANDARD; PRT; 72 AA.  
ID YF77\_HAEIN  
AC Q57070; P96345;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL PROTEIN HI1577.  
GN HI1577.

OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McElroy K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus  
RT influenzae Rd."  
RL Science 269:496-512(1995).

CC -1- SIMILARITY: STRONG, TO H. INFLUENZAE HI1329.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U32832; AAC23226.1; -  
DR TIGR: HI1577; -  
KW Hypothetical protein.  
SQ SEQUENCE 72 AA; 8370 MW; D00DAB6A8E96660 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 VNINKL 541  
Db 31 VNINKL 36  
|||||

RESULT 134  
Y055\_TREPA STANDARD; PRT; 78 AA.  
ID Y055\_TREPA  
AC O83094;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE HYPOTHETICAL PROTEIN TP0055.  
GN TP0055.

OS Treponema pallidum.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
OX NCBI\_TaxID=160;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=NICHOLS;  
RX MEDLINE=98332770; PubMed=9665876;

RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,  
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
RA Venter J.C.;

RT "Complete genome sequence of Treponema pallidum, the syphilis  
RT spirochete."  
RL Science 281:375-388(1998).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AE001190; AAC65053.1; -  
DR TIGR: TP0055; -  
KW Hypothetical protein; Transmembrane  
FT TRANSMEM 12 32 POTENTIAL.  
FT TRANSMEM 51 71 POTENTIAL.  
SQ SEQUENCE 78 AA; 7934 MW; AEDCD20DBDBAE5A4 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 78;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VGLVLS 25  
Db 35 VGLVLS 40  
|||||

RESULT 135  
MOAD\_HAEIN STANDARD; PRT; 81 AA.  
ID MOAD\_HAEIN  
AC P45309;  
DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE MOLYBDOPTERIN [MPT] CONVERTING FACTOR, SUBUNIT 1 (MOLYBDENUM COFACTOR  
DE BIOSYNTHESIS PROTEIN D) (MOLYBDOPTERIN CONVERTING FACTOR SMALL  
DE SUBUNIT).  
GN MOAD OR HI1674.

OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;

```

RN SEQUENCE FROM N.A.
RP STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE-95350630; PubMed-7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.;"
RL Science 269:496-512(1995).
[2]
RN REVISIONS.
RA White O., Kerlavage A.R., Fleischmann R.D.;
RA Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CONVERSION OF MOLYBDOTERIN PRECURSOR 2 INTO
CC MOLYBDOTERIN REQUIRES TRANSFER OF TWO SULFUR ATOMS TO PRECURSOR 2
CC (TO GENERATE THE DITHIOLENE GROUP). THIS IS CATALYZED BY THE
CC CONVERTING FACTOR COMPOSED OF A SMALL AND LARGE SUBUNIT. THE
CC SULFUR ATOMS ARE PROVIDED BY THE ACTIVE FORM OF THE SMALL SUBUNIT,
CC WHOSE ACTIVATION INVOLVES THE ACQUISITION OF SULFUR AND THE
CC ACTIVITY OF MOEB/CHLN (BY SIMILARITY).
CC -!- PATHWAY: MOLYBDENUM COFACTOR BIOSYNTHESIS.
CC -!- SUBUNIT: HETERODIMER OF MOAD AND MOAE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32840; AAC23319.1; -
DR TIGR; H11674; -
KW Molybdenum cofactor biosynthesis.
SQ SEQUENCE 81 AA; 8826 MW; 35D1440F82456F22 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 878 LVAINQ 883
DB 52 LVAINQ 57
|||||
PRT; 81 AA.
STANDARD; PRT;
ID VES_HPV35 STANDARD; PRT;
AC P27226;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROBABLE E5 PROTEIN.
GN E5.
OS Human papillomavirus type 35.
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
OX NCBI_TaxID=10587;
[1]
RN SEQUENCE FROM N.A.
RA Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of
RT human papillomavirus type 35.;"
RL Virology 186:770-776(1992).
-----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74117; AAA46970.1; -
DR PIR; D40824; W5WL35.
KW Early protein.
SQ SEQUENCE 81 AA; 9000 MW; B12A23102E72163B CRC64;

Query Match 0.5%; Score 6; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ASSTVL 220
DB 6 ASSTVL 11
|||||
PRT; 85 AA.
STANDARD; PRT;
ID YOR4_BPSP STANDARD; PRT;
AC Q38440;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL 10.2 KDA PROTEIN IN GP2-GP6 INTERGENIC REGION (ORF 4).
OS Bacteriophage SPPI.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OX Lambda phage group.
OX NCBI_TaxID=10724;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-92194332; PubMed=1548711;
RA Chai S., Bravo A., Lueder G., Nedlin A., Trautner T.A., Alonso J.C.;
RT "Molecular analysis of the Bacillus subtilis bacteriophage SPPI
RT region encompassing genes 1 to 6. The products of gene 1 and gene 2
RT are required for pac cleavage.;"
RL J. Mol. Biol. 224:87-102(1992).
[2]
RN SEQUENCE FROM N.A.
RP Alonso J.C., Luder G., Stiege A.C., Chai S., Weise F., Trautner T.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: SOME, TO PHAGE SP01 GENE 46 PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56064; CAA39539.1; -
DR EMBL; X97918; CAA66577.1; -
KW Hypothetical protein.
SQ SEQUENCE 85 AA; 10206 MW; 73FABC4D8CE2DF86 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1234 AASLNT 1239
DB 48 AASLNT 53
|||||
PRT; 53
STANDARD; PRT;
ID AASLNT 1239
AC AASLNT 53
DT 12-APR-1992 (Rel. 23, Created)
DT 12-APR-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROBABLE E5 PROTEIN.
GN E5.
OS Human papillomavirus type 35.
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
OX NCBI_TaxID=10587;
[1]
RN SEQUENCE FROM N.A.
RA Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of
RT human papillomavirus type 35.;"
RL Virology 186:770-776(1992).
-----

```

RESULT 138

MOAA\_METTM STANDARD; PRT; 87 AA.  
AC Q50746;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE PUTATIVE MOLYBDOPROTEIN COFACTOR SYNTHESIS PROTEIN A (FRAGMENT).  
GN MOAA.  
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanothermobacter.  
OX NCBI\_TaxID=79929;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96163477; PubMed=8575452;  
RA Hochheimer A., Schmitz R.A., Thauer R.K., Hedderich R.;  
RT "The tungsten formylmethanofuran dehydrogenase from Methanobacterium  
RT thermoautotrophicum contains sequence motifs characteristic for  
RT enzymes containing molybdopterin dinucleotide.";  
RL Eur. J. Biochem. 234:910-920(1995).  
CC -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF MOLYBDOPROTEIN PRECURSOR  
CC Z FROM GUANOSINE (BY SIMILARITY).  
CC -!- PATHWAY: MOLYBDENUM COFACTOR BIOSYNTHESIS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE MOAA / NIFB / POQE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X87968; CAA61207.1; -;  
DR InterPro; IPR000345; -;  
DR InterPro; IPR000385; -;  
DR PROSITE; PS01305; MOAA\_NIFB\_POQE; 1.  
KW Molybdenum cofactor biosynthesis; Iron-sulfur.  
FT METAL 22 22 IRON-SULFUR (POTENTIAL).  
FT METAL 26 26 IRON-SULFUR (POTENTIAL).  
FT METAL 29 29 IRON-SULFUR (POTENTIAL).  
FT NON\_TER 87 87  
SQ SEQUENCE 87 AA; 9874 MW; A97432C6A439F909 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 87;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 12 RPLVSL 17  
DB 9 RPLVSL 14  
  
RESULT 139  
YLS9\_CAEEL STANDARD; PRT; 91 AA.  
ID YLS9\_CAEEL  
AC P34394;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE HYPOTHETICAL 11.1 KDA PROTEIN F09G8.9 IN CHROMOSOME III.  
GN F09G8.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,  
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Watson R., Watson A., Weinstein L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT \*2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.;  
RL Nature 368:32-38(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L11247; AAA28003.1; -;  
DR PIR; S44791; S44791.  
DR WormPep; F09G8.9; CE00144.  
KW Hypothetical protein.  
SQ SEQUENCE 91 AA; 11141 MW; 59F7D3CECDC1D400 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 91;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 805 NYKYLI 810  
DB 17 NYKYLI 22  
  
RESULT 140  
VP3\_SSV1 STANDARD; PRT; 92 AA.  
ID VP3\_SSV1  
AC P20225;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE STRUCTURAL PROTEIN VP3.  
OS Sulfolobus virus-like particle SSV1.  
OC Viruses; dsDNA viruses, no RNA stage; Fuselloviridae; Fusellovirus.  
OX NCBI\_TaxID=10476;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-4.  
RA Reiter W.-D., Palm P., Henschen A., Lottspeich F., Zillig W.,  
RA Grampp B.;  
RT "Identification and characterization of the genes encoding three  
RT structural proteins of the Sulfolobus virus-like particle SSV1.";  
RL Mol. Gen. Genet. 206:144-153(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9204080; PubMed=1926776;  
RA Palm P., Schleper C., Grampp B., Yeats S., McWilliam P., Reiter W.-D.,  
RA Zillig W.;  
RT "Complete nucleotide sequence of the virus SSV1 of the  
RT archaeobacterium Sulfolobus shibatae.";  
RL Virology 185:242-250(1991).  
CC -!- SIMILARITY: TO SSV1 VP1.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```

DR EMBL; X04829; CAA28516.1; -
DR EMBL; X07234; CAA30203.1; -
DR PIR; C40782; VCDX3.
KW Structural protein; Transmembrane.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
SQ SEQUENCE 92 AA; 9868 MW; 8C796F374B5DF24E CRC64;

Query Match 0.5%; Score 6; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1229 VGSNNA 1234
DB 58 VGSNNA 63

RESULT 141
CH10_BORBU STANDARD; PRT; 93 AA.
AC O51683;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 10 KDA CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
GN MOPB OR GROES OR BB0741.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RA MEDLINE-98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kertavagis A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Wathley L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -!- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
CC THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).
CC -!- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001174; AAC67092.1; -
DR TIGR; BB0741; -
DR InterPro; IPR001476; -
DR Pfam; PF00166; cpn10; 1.
DR PRINTS; PR00297; CHAPERONIN10.
DR PROSITE; PS00681; CHAPERONINS_CPN10; 1.
KW Chaperone.
SQ SEQUENCE 93 AA; 10313 MW; E9281F76CAD94E4B CRC64;

Query Match 0.5%; Score 6; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 307 KTNIGT 312
DB 38 KTNIGT 43

RESULT 142
REPI_ECOLI STANDARD; PRT; 93 AA.
AC P05830;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE REGULATORY PROTEIN REPI (FRAGMENT).
GN REPI.
OS Escherichia coli.
OG Plasmid R1162.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87040771; PubMed=2430262;
RA Kim K., Meyer R.J.;
RT "Copy-number of broad host-range plasmid R1162 is regulated by a
RT small RNA.";
RL Nucleic Acids Res. 14:8027-8046(1986).
CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN REGULATING THE PLASMID
CC COPY-NUMBER. INCREASING THE LEVEL OF THIS PROTEIN RESULTS IN A
CC HIGHER PLASMID COPY-NUMBER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04499; CAA28184.1; -
DR PIR; A25546; RGECRI.
KW Plasmid; Plasmid copy control.
FT NON_TER 93
SQ SEQUENCE 93 AA; 9413 MW; 0E767AAC26C7634A CRC64;

Query Match 0.5%; Score 6; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VGLVLS 25
DB 32 VGLVLS 37

RESULT 143
ESA6_MYCLE STANDARD; PRT; 95 AA.
AC Q50206; Q33083;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 6 KDA EARLY SECRETORY ANTIGENIC TARGET HOMOLOG (ESAT-6-LIKE PROTEIN)
DE (L-ESAT).
GN ESAT6 OR ESX OR L45 OR MLCB628.12C.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Wieles B., Notenboom T., Naafs B., Offringa R., Ottenhoff T.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]

```



```

RP SEQUENCE FROM N.A.
RA Eglmeier K., Garnier T., De Rossi E., Psihi H., Cole S.T.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X90946; CAA62441.1; -
DR EMBL; Y14967; CAA75200.1; -
FT CONFLICT 55 55 R -> Q (IN REF. 1).
FT CONFLICT 90 90 M -> T (IN REF. 1).
SQ SEQUENCE 95 AA; 10465 MW; B1526F78CB2AB8A1 CRC64;
-----
Query Match 0.5%; Score 6; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 578 GSQSR1 583
Db .19 GSQSR1 24
|||||

RESULT 144
Y400_HAEIN STANDARD; PRT; 95 AA.
AC P44686;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL PROTEIN HI0400.
GN HI0400.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae RD.";
RT Science 269:496-512(1995).
CC -!- SIMILARITY: BELONGS TO THE YFCZ/YIIS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U3723; AAC22059.1; -
DR TIGR; HI0400; -
KW Hypothetical protein.
SQ SEQUENCE 95 AA; 10342 MW; 1B09AB38912AA953 CRC64;

```

```

Query Match 0.5%; Score 6; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 893 FELANR 898
Db 90 FELANR 95
|||||

RESULT 145
YQEI_BACSU STANDARD; PRT; 96 AA.
AC P54454;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL 10.8 KDA PROTEIN IN AROD-COMER INTERGENIC REGION.
GN YQEI.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Kobayashi Y., Mizuno M., Masuda S., Takemaru K., Hosono S.,
RA Sato T., Takeuchi M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE UPF0044 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D84432; BAA12446.1; -
DR EMBL; Z99117; CAB14507.1; -
DR Subtilist; BG11637; yqei.
DR InterPro; IPR001890; -
DR Pfam; PF01985; UPF0044; 1.
DR PROSITE; PS01301; UPF0044; 1.
KW Hypothetical protein.
SQ SEQUENCE 96 AA; 10758 MW; D9BA3E935C161F54 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 430 GKGGVN 435
Db 24 GKGGVN 29
|||||

RESULT 146
MOBS_THIFE STANDARD; PRT; 98 AA.
AC P20086;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE MOBILIZATION PROTEIN MOBS.
GN MOBS.
OS Thiobacillus ferrooxidans.
OG Plasmid pTFL.
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
OX NCBI_TaxID=920;
[1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-ATCC 33020;  
RX MEDLINE-91125140; PubMed-2280689;  
RA Drolet M., Zanga P., Lau P.C.K.;  
RT "The mobilization and origin of transfer regions of a Thiobacillus  
ferrooxidans plasmid: relatedness to plasmids RSP1010 and pSC101.";  
RL Mol. Microbiol. 4:1381-1391(1990).  
CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO PROMOTE THE SPECIFIC  
TRANSFER OF THE PLASMID IN THE PRESENCE OF CONJUGATIVE PLASMIDS.  
CC -!- SIMILARITY: TO MOBILIZATION PROTEIN C OF THE E.COLI PLASMID  
INCO RSP1010.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X52699; CAA36926.1; -.  
DR PIR; S12189; S12189.  
KW Mobility protein; Plasmid; Trans-acting factor; Conjugation.  
SQ SEQUENCE 98 AA; 11451 MW; 536785B47C9E3F54 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 19 LVGALV 24  
Db ||||||  
54 LVGALV 59  
  
RESULT 147  
SECY\_BACST STANDARD; PRT; 99 AA.  
AC P28620;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE PREPROTEIN TRANSLOCASE SECY SUBUNIT (FRAGMENT).  
GN SECY.  
OS Bacillus stearothermophilus.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1422;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92207915; PubMed-1554691;  
RA Glaser P., Presecan E., Delepiere M., Surewicz W.K.,  
RA Mantsch H.H., Barzu O., Gilles A.M.;  
RT "Zinc, a novel structural element found in the family of bacterial  
adenylate kinases.";  
RL Biochemistry 31:3038-3043(1992).  
CC -!- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH SECA AND SECE  
TO ALLOW THE TRANSLOCATION OF PROTEINS ACROSS THE PLASMA MEMBRANE,  
BY FORMING PART OF A CHANNEL.  
CC -!- SUBUNIT: ONE OF SEVEN SECRETORY PROTEINS (SECA-F & SECY) THAT  
COMPRISE THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE SECY/SEC61-ALPHA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M88104; AAA22204.1; -.  
DR PIR; A42196; A42196.

DR InterPro: IPR002208; -.  
DR Pfam: PF00344; secY; 1.  
DR PROSITE; PS00755; SECY\_1; PARTIAL.  
DR PROSITE; PS00756; SECY\_2; PARTIAL.  
KW Protein transport; Translocation; Transmembrane.  
FT NON\_TER 1 1  
FT DOMAIN <1 35 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 36 55 POTENTIAL.  
FT DOMAIN 56 60 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 61 79 POTENTIAL.  
FT DOMAIN 80 99 CYTOPLASMIC (POTENTIAL).  
SQ SEQUENCE 99 AA; 10910 MW; F33BB10867C7C88B CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1035 IGGTSL 1040  
Db ||||||  
66 IGGTSL 71  
  
RESULT 148  
TF1\_BPSP1 STANDARD; PRT; 99 AA.  
AC P04445;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE TRANSCRIPTION FACTOR 1.  
GN TF1.  
OS Bacteriophage SP01.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.  
OC SP01-like Viruses.  
OX NCBI\_TaxID=10685;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE-85063726; PubMed-6438630;  
RA Greene J.R., Brennan S.M., Andrew D.J., Thompson C.C., Richards S.H.,  
RA Heinrichson R.L., Geiduschek E.P.;  
RT "Sequence of the bacteriophage SP01 gene coding for transcription  
factor 1, a viral homologue of the bacterial type II DNA-binding  
proteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:7031-7035(1984).  
RN [2]  
RP MUTAGENESIS.  
RX MEDLINE-91094044; PubMed-2125081;  
RA Sayre M.H., Geiduschek E.P.;  
RT "Effects of mutations at amino acid 61 in the arm of TF1 on its DNA-  
binding properties.";  
RL J. Mol. Biol. 216:819-833(1990).  
RN [3]  
RP STRUCTURE BY NMR.  
RX MEDLINE-93238776; PubMed-8477755;  
RA Reisman J.M., Hsu V.L., Jarriel-Encontre I., Lecou C., Sayre M.H.;  
RA Kearns D.R., Parello J.;  
RT "A 1H-NMR study of the transcription factor 1 from Bacillus subtilis  
phage SP01 by selective 2H-labeling. Complete assignment and  
structural analysis of the aromatic resonances for a 22-kDa  
homodimer.";  
RL Eur. J. Biochem. 213:865-873(1993).  
RN [4]  
RP STRUCTURE BY NMR.  
RX MEDLINE-97070379; PubMed-8913305;  
RA Jia X., Grove A., Ivancic M., Hsu V.L., Geiduschek E.P., Kearns D.R.;  
RT "Structure of the Bacillus subtilis phage SP01-encoded type II DNA-  
binding protein TF1 in solution.";  
RL J. Mol. Biol. 263:259-268(1996).  
CC -!- FUNCTION: TF1 SELECTIVELY BINDS TO AND INHIBITS THE TRANSCRIPTION  
OF HYDROXYMETHYLURACIL-(HM-URA)-CONTAINING DNA, SUCH AS SP01 DNA,  
BY RNA POLYMERASE IN VITRO.  
CC -!- SUBUNIT: HOMODIMER.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL HISTONE-LIKE PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; K02381; AAA32599.1; -  
DR PIR; A02693; DNBPI1.  
DR PIR; S13254; S13254.  
DR PDB; 1WTU; 12-FEB-97.  
DR InterPro; IPR000119; -  
DR Pfam; PF00216; Bac\_DNA\_Binding; 1.  
DR PROSITE; PS00045; HISTONE\_LIKE; 1.  
KW Transcription regulation; Repressor; DNA-binding; 3D-structure.  
FT DNA\_BIND 61  
FT DNA\_BIND 52 77 MAY BE INVOLVED IN PREFERENCE FOR HM-URA  
FT DOMAIN 90 99 MAY BE INVOLVED IN PREFERENCE FOR HM-URA  
FT P37854; -  
FT DNA\_BIND 93 94 POTENTIAL.  
FT SEQUENCE 99 AA; 10737 MW; 3F8CD237A46EFC37 CRC64;  
SQ

Query Match 0.5%; Score 6; DB 1; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1164 PSVGVS 1169  
Db 73 PSVGVS 78

RESULT 149  
CHLB\_PLESC STANDARD; PRT; 100 AA.  
AC P37854;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE LIGHT-INDEPENDENT PROTOCHLOROPHYLLIDE REDUCTASE CHLB SUBUNIT  
DE (EC 1.3.1.33) (DPOR CHLB SUBUNIT) (FRAGMENT).  
GN CHLB.  
OS Pleurozium schreberi (Moss).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Bryopsida; Bryidae;  
OC Hypnales; Hylocomiaceae; Pleurozium.  
OX NCBI\_TaxID=34163;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Boivin R., Richard M., Bousquet J., Bellemare G.;  
RL Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: INVOLVED IN THE LIGHT-INDEPENDENT ACCUMULATION OF  
CC CHLOROPHYLL, PROBABLY AT THE STEP OF REDUCTION OF PROTO-  
CC CHLOROPHYLLIDE TO CHLOROPHYLLIDE (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: CHLOROPHYLLIDE A + NADP(+) =  
CC PROTOCHLOROPHYLLIDE + NADPH.  
CC -1- PATHWAY: LIGHT-INDEPENDENT CHLOROPHYLL BIOSYNTHESIS.  
CC -1- SIMILARITY: BELONGS TO THE CHLB / BCHB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L25774; AAC37495.1; -  
DR HSSP; P11347; LM10.

DR Mendel; 2284; PLESC; chlb; 1.  
DR InterPro; IPR000510; -  
DR Pfam; PF00148; oxidored\_nitro; 1.  
KW Chloroplast; Photosynthesis; Chlorophyll biosynthesis; Oxidoreductase;  
KW NADP.  
FT NON\_TER 1  
FT NON\_TER 100 100  
FT SEQUENCE 100 AA; 11659 MW; 40B101BE6D66B5E0 CRC64;  
SQ

Query Match 0.5%; Score 6; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1128 LLODLN 1133  
Db 3 LLODLN 8

RESULT 150  
CHLB\_POLCU STANDARD; PRT; 100 AA.  
AC P37852;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE LIGHT-INDEPENDENT PROTOCHLOROPHYLLIDE REDUCTASE CHLB SUBUNIT  
DE (EC 1.3.1.33) (DPOR CHLB SUBUNIT) (FRAGMENT).  
GN CHLB.  
OS Polytrichum commune (Moss).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Polytrichopsida;  
OC Polytrichales; Polytrichaceae; Polytrichum.  
OX NCBI\_TaxID=3213;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Boivin R., Richard M., Bousquet J., Bellemare G.;  
RL Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: INVOLVED IN THE LIGHT-INDEPENDENT ACCUMULATION OF  
CC CHLOROPHYLL, PROBABLY AT THE STEP OF REDUCTION OF PROTO-  
CC CHLOROPHYLLIDE TO CHLOROPHYLLIDE (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: CHLOROPHYLLIDE A + NADP(+) =  
CC PROTOCHLOROPHYLLIDE + NADPH.  
CC -1- PATHWAY: LIGHT-INDEPENDENT CHLOROPHYLL BIOSYNTHESIS.  
CC -1- SIMILARITY: BELONGS TO THE CHLB / BCHB FAMILY.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L25772; AAC37494.1; -  
CC Mendel; 2285; POLCU; chlb; 1.  
CC InterPro; IPR000510; -  
CC Pfam; PF00148; oxidored\_nitro; 1.  
KW Chloroplast; Photosynthesis; Chlorophyll biosynthesis; Oxidoreductase;  
KW NADP.  
FT NON\_TER 1  
FT NON\_TER 100 100  
FT SEQUENCE 100 AA; 11573 MW; 984F131F432400A3 CRC64;  
SQ

Query Match 0.5%; Score 6; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1128 LLODLN 1133  
Db 3 LLODLN 8

RESULT 151  
GATC\_RICPR STANDARD; PRT; 100 AA.  
AC G9ZE09;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT C (EC 6.3.5.-) (GLU-ADT  
SUBUNIT C).  
GN GATC OR RP153.  
OS Rickettsia prowazekii.  
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
OX NCBI\_TaxID=782;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MADRID E;  
RX MEDLINE=99039459; PubMed=9823893;  
RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,  
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
RT "The genome sequence of Rickettsia prowazekii and the origin of  
RT mitochondria";  
RL Nature 396:133-140(1998).  
CC -!- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED  
CC GLN-TRNA(GLN) THROUGH THE TRANSAMIDATION OF MISACYLATED GLU-  
CC TRNA(GLN) IN ORGANISMS WHICH LACK GLUTAMINYL-TRNA SYNTHETASE. THE  
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH  
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-TRNA(GLN) (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMYL-TRNA(GLN) + L-GLUTAMINE = ADP  
CC + PHOSPHATE + L-GLUTAMINYL-TRNA(GLN) + L-GLUTAMATE.  
CC -!- SUBUNIT: HETEROTRIMER OF A, B AND C SUBUNITS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE GATC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AJ235270; CAA14621.1; -  
DR PIR: G43258; G43258.  
DR HSP: P23532; IE2A.  
KW Phosphotransferase system; Sugar transport; Transferase;  
FT MOD\_RES 78 78 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 100 AA; 11344 MW; 9640F6A44A685474 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1262 KLAKEV 1267  
Db 85 KLAKEV 90  
|||||  
RESULT 152  
PTLA\_STRMU STANDARD; PRT; 104 AA.  
AC P26426;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE PTS SYSTEM, LACTOSE-SPECIFIC IIA COMPONENT (EIIA-LAC) (LACTOSE-  
DE PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT)  
DE (EC 2.7.1.69) (EIIA-LAC).  
GN LACF.  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus  
OX NCBI\_TaxID=1309;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=93015655; PubMed=1400164;  
RA Rosey E.L., Stewart G.C.;  
RT "Nucleotide and deduced amino acid sequences of the lacR, lacABCD,  
RT and lacFE genes encoding the repressor, tagatose 6-phosphate gene  
RT cluster, and sugar-specific phosphotransferase system components of  
RT the lactose operon of Streptococcus mutans.";  
RL J. Bacteriol. 174:6159-6170(1992).  
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT  
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE  
CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY  
CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS  
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO  
CC THE SUGAR.  
CC -!- CATALYTIC ACTIVITY: PROTEIN N-PHOSPHOHISTIDINE + SUGAR =  
CC PROTEIN HISTIDINE + SUGAR PHOSPHATE.  
CC -!- SUBUNIT: HOMOTRIMER.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: CONTAINS A PTS EIIA DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M80797; AAA26908.1; -  
DR PIR: S27705; S27705.  
DR PIR: G43258; G43258.  
DR HSP: P23532; IE2A.  
KW Phosphotransferase system; Sugar transport; Transferase;  
FT MOD\_RES 78 78 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 104 AA; 11399 MW; 281CB2F3CB109F5D CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 104;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 296 NAAQAG 301  
Db 29 NAAQAG 34  
|||||  
RESULT 153  
FUMH\_MOUSE STANDARD; PRT; 105 AA.  
AC P97807;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE FUMARATE HYDRATASE, MITOCHONDRIAL (EC 4.2.1.2) (FUMARASE) (EF-3)  
DE (FRAGMENT).  
GN FH OR FHL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97184476; PubMed=9032278;  
RA Fu X., Kamps M.P.;  
RT "F2a-Pbx1 induces aberrant expression of tissue-specific and  
RT developmentally regulated genes when expressed in NIH 3T3  
RT fibroblasts";  
RL Mol. Cell. Biol. 17:1503-1512(1997).  
CC -!- CATALYTIC ACTIVITY: L-MALATE = FUMARATE + H(2)O.  
CC -!- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL AND CYTOPLASMIC.  
CC -1- SIMILARITY: TO OTHER THERMOSTABLE CLASS II FUMARASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U72679; AA851039.1; -  
DR HSP; P08417; IYFM.  
DR MGD; MGI:95530; Fhl.  
DR InterPro: IPR000362; -  
DR Pfam: PF00206; lyase\_1; 1.  
DR PROSITE; PS00163; FUMARATE\_LYASES; PARTIAL.  
KW Lyase; Tricarboxylic acid cycle; Mitochondrion.  
FT NON\_TER 1 1  
FT ACT\_SITE 55 55 POTENTIAL.  
FT NON\_TER 105 105  
FT NON\_TER 105 105  
SQ SEQUENCE 105 AA; 11233 MW; 4F7E9DBCED39456A CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 105;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 51 GRAVGT 56  
Db 96 GRAVGT 101  
|||||  
  
RESULT 154  
COL\_RABBIT STANDARD; PRT; 107 AA.  
ID COL\_RABBIT  
AC P42890;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE COLIPASE PRECURSOR.  
GN CLPS.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RX MEDLINE=93345715; PubMed=8344444;  
RA Colwell N.S., Aleman-Gomez J.A., Sasser T.L., Kumar V.B.;  
RT "Cloning and characterization of rabbit pancreatic colipase."  
RL Int. J. Biochem. 25:885-890(1993).  
CC -1- FUNCTION: COLIPASE IS A COFACTOR OF PANCREATIC LIPASE. IT ALLOWS  
CC THE LIPASE TO ANCHOR ITSELF TO THE LIPID-WATER INTERFACE. WITHOUT  
CC COLIPASE THE ENZYME IS WASHED OFF BY BILE SALTS, WHICH HAVE AN  
CC INHIBITORY EFFECT ON THE LIPASE.  
CC -1- FUNCTION: ENTEROSTATIN HAS A BIOLOGICAL ACTIVITY AS A SATIETY  
CC SIGNAL (BY SIMILARITY).  
CC -1- SUBUNIT: FORM A 1:1 STOICHIOMETRIC COMPLEX WITH PANCREATIC LIPASE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L06329; AAA02911.1; -  
DR HSP; P02703; LPCO.  
DR InterPro: IPR001981; -  
DR Pfam; PF01114; Colipase; 1.

DR PRINTS; PRO0128; COLIPASE.  
DR PROSITE; PS00121; COLIPASE; 1.  
KW Lipid degradation; Digestion; Pancreas; Signal.  
FT SIGNAL 1 17 POTENTIAL.  
FT PROPEP 18 22 ENTEROSTATIN, ACTIVATION PEPTIDE  
(POTENTIAL).  
FT CHAIN 23 107 COLIPASE.  
FT DISULFID 34 45 BY SIMILARITY.  
FT DISULFID 40 56 BY SIMILARITY.  
FT DISULFID 44 78 BY SIMILARITY.  
FT DISULFID 66 86 BY SIMILARITY.  
FT DISULFID 80 104 BY SIMILARITY.  
SQ SEQUENCE 107 AA; 11271 MW; 825BA1AFB1422390 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 107;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1091 NTNFGV 1096  
Db 98 NTNFGV 103  
|||||  
  
RESULT 155  
YNIU\_AZOV STANDARD; PRT; 107 AA.  
ID YNIU\_AZOV  
AC Q44540;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL 11.0 KDA PROTEIN IN NIFU 5'REGION (ORF6).  
OS Azotobacter vinelandii.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Azotobacter.  
OX NCBI\_TaxID=354;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89123097; PubMed=2644218;  
RA Jacobson M.R., Brigle K.E., Bennett L.T., Setterquist R.A.,  
RA Wilson M.S., Cash V.L., Beynon J., Newton W.E., Dean D.R.;  
RT "Physical and genetic map of the major nif gene cluster from  
RT Azotobacter vinelandii."  
RL J. Bacteriol. 171:1017-1027(1989).  
CC -1- SIMILARITY: BELONGS TO THE HESB/YADR/YFHF FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M20568; AAA64724.1; -  
DR InterPro: IPR000361; -  
DR Pfam; PF01521; HESB-like; 1.  
DR PROSITE; PS01152; HESB; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 107 AA; 11047 MW; 27BE5FE151447219 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 107;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 919 TLLIDS 924  
Db 61 TLLIDS 66  
|||||  
  
RESULT 156  
YTXJ\_BACSU

```

ID YTXJ_BACSU STANDARD; PRT; 108 AA.
AC P39914;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHEICAL 12.4 KDA PROTEIN IN MURC-AROA INTERGENIC REGION (ORF2)
DE (ORF3).
GN YTXJ.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Bolotin A.P., Khazak V.E., Ratmanova K.I., Yomantas Y.I.,
RA Kozlov Y.I.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb rrrB-dnaB region.";
RL Microbiology 143:3431-3441(1997).
RN [3]
RP SEQUENCE OF 1-56 FROM N.A.
RC STRAIN=168 / MARBURG;
RX MEDLINE=96310371; PubMed=8733232;
RA Varon D., Brody M.S., Price C.W.;
RT "Bacillus subtilis operon under the dual control of the general
RT stress transcription factor sigma B and the sporulation transcription
RT factor sigma H.";
RL Mol. Microbiol. 20:339-350(1996).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X65945; CAA46762.1; -
DR EMBL; AF008220; AAC00297.1; -
DR EMBL; L31845; AAB40046.1; -
DR EMBL; Z99119; CAB14954.1; -
DR PIR; S21420; S21420.
DR Subtilist; BG10373; ytxJ.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 12402 MW; D8F03DD68E79AF38 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1101 FANQHE 1106
DB 44 FANQHE 49

RESULT 157
ID VHEM_PVMR STANDARD; PRT; 109 AA.
AC P17527;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE 12 KDA MEMBRANE PROTEIN (ORF 3).
OS Potato virus M (strain Russian) (PVM).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
OX NCBI_TaxID=12168;


```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89293091; PubMed=2738581;
RA Rupasov V.V., Morozov S.Y., Kanyuka K.V., Zavriev S.K.;
RT "Partial nucleotide sequence of potato virus M RNA shows similarities
RT to protoviruses in gene arrangement and the encoded amino acid
RT sequences.";
RL J. Gen. Virol. 70:1861-1869(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91116326; PubMed=1990070;
RA Zavriev S.K., Kanyuka K.V., Levay K.E.;
RT "The genome organization of potato virus M RNA.";
RL J. Gen. Virol. 72:9-14(1991).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: TO ORF3 PROTEIN FROM POTEXVIRUSES AND TO THE 14 KDA
CC PROTEIN FROM BSMV RNA 2BETA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14449; BAA03341.1; -
DR EMBL; X53062; CAA37234.1; -
DR PIR; PN0003; WMYP2.
DR PIR; S21603; S21603.
DR PIR; C54333; C54333.
DR InterPro; IPR001896; -
DR Pfam; PF01307; Plant_vir_prot; 1.
KW Transmembrane.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
SQ SEQUENCE 109 AA; 11907 MW; F31D213AA67CD1AA CRC64;

Query Match 0.5%; Score 6; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VSLALV 20
DB 20 VSLALV 25

RESULT 158
ID VG21_BPMD2 STANDARD; PRT; 111 AA.
AC O64215;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENE 21 PROTEIN (GP21).
GN 21.
OS Mycobacteriophage D29.
OC Viruses.
OX NCBI_TaxID=28369;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: implications for phage
RT evolution.";
RL J. Mol. Biol. 279:143-164(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC -----

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC -----
DR EMBL: AF022214; AAC18462.1; -.
SQ SEQUENCE 111 AA; 12155 MW; 9E51DE732E2232AF CRC64;
```

Query Match	0.5%;	Score 6;	DB 1;	Length 111;
Best Local Similarity	100.0%			
Best Ne	2	1	100	

```

Best local similarity 100.0%, PRed. NO. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 482 AGDTK 487

db 88 AGTDTK 93

RESULT 159

VG21_BPML5	ID	VG21_BPML5	STANDARD;	PRT;	111 AA;
------------	----	------------	-----------	------	---------

AC Q05227;  
DT 01-FEB-1994 (Rel. 28, Created)

DT	Rel. 28, created)
DT	01-FEB-1994 (Rel. 28, Last sequence update)
DT	01-FEB-1994 (Rel. 28, Last sequence update)

01-FEB-1994 (REL. 28, last annotation update)  
DE GENE 21 PROTEIN (GP21).

```

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=UAL 187;
RX MEDLINE=92041660; PubMed=1840587;
RA Hastings J.W., Sailer M., Johnson K., Roy K.L., Vederas J.C.,
RA Stiles M.E.;
RT "Characterization of leucocin A-UAL 187 and cloning of the
RL bacteriocin gene from Leuconostoc gelidum.";
RL J. Bacteriol. 173:7491-7500(1991).
CC -1- FUNCTION: IMPARTS IMMUNITY TO LEUCOCIN A TO NATURALLY SENSITIVE
CC HOST STRAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M64371; AAA68004.1; -
DR PIR; B41657; B41657.
KW Bacteriocin immunity; Plasmid.
SQ SEQUENCE 113 AA; 12947 MW; B83B5C83F30C09D4 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 542 ITASTN 547
DB 64 ITASTN 69

RESULT 162
MIF_CHICK
ID MIF_CHICK STANDARD; PRT; 114 AA.
AC Q02960;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MACROPHAGE MIGRATION INHIBITORY FACTOR (MIF).
GN MIF.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN SEQUENCE FROM N.A.
RP TISSUE=Lens;
RX MEDLINE=93165679; PubMed=7679497;
RA Wistow G.J., Shaugnessy M., Lee D.C., Hodin J., Zelenka P.S.;
RT "A macrophage migration inhibitory factor is expressed in the
RL differentiating cells of the eye lens.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:1272-1275(1993).
CC -1- FUNCTION: THE EXPRESSION OF MIF AT SITES OF INFLAMMATION SUGGEST
CC A ROLE FOR THE MEDIATOR IN REGULATING THE FUNCTION OF MACROPHAGE
CC IN HOST DEFENSE.
CC -----
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MIF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95776; AAA48939.1; -
DR PIR; C47274; C47274.

```

```

DR HSSP; P14174; IGIF.
DR InterPro; IPR001398; -
DR Pfam; PF01187; MIF; 1.
DR PROSITE; PS01158; MIF; 1.
KW Macrophage; Inflammatory response; Cytokine.
FT INIT_MET 0 BY SIMILARITY.
SQ SEQUENCE 114 AA; 12353 MW; A55222D00E6D05CF CRC64;

Query Match 0.5%; Score 6; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 GGOQNK 131
DB 68 GGOQNK 73

RESULT 163
NLTI_LYCES
ID NLTI_LYCES STANDARD; PRT; 114 AA.
AC P93224;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NONSPECIFIC LIPID-TRANSFER PROTEIN 1.PRECURSOR (LTP 1).
GN LEL6.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN SEQUENCE FROM N.A.
RP STRAIN=CV. AITLS CRAIG;
RA Plant A.L., Cohen A., Bray E.A.;
RT "Nucleotide sequence and spatial expression pattern of a drought- and
RL abscisic acid-induced gene of tomato.";
RL Plant Physiol. 97:900-906(1991).
CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES.
CC -1- INDUCTION: BY DROUGHT AND ABA.
CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U81996; AAB42069.1; -
DR HSSP; P19656; 1AFH.
DR InterPro; IPR000528; -
DR Pfam; PF00279; LTP; 1.
DR PRINTS; PR00382; LIPIDTRNSFR.
DR PROSITE; PS00597; PLANT_LTP; 1.
KW Lipid-binding; Transport; Signal; Multigene family..
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 114 NONSPECIFIC LIPID-TRANSFER PROTEIN 1.
FT DISULFID 27 73 BY SIMILARITY.
FT DISULFID 37 50 BY SIMILARITY.
FT DISULFID 51 96 BY SIMILARITY.
FT DISULFID 71 110 BY SIMILARITY.
SQ SEQUENCE 114 AA; 11484 MW; DB4989CF02E39FDA CRC64;

Query Match 0.5%; Score 6; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

QY 515 KGIDTG 520
Db 82 KGIDTG 87

RESULT 164
SODC_DROOB
ID SODC_DROOB STANDARD; PRT; 114 AA.
AC Q95085;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SUPEROXIDE DISMUTASE [CU-ZN] (EC 1.15.1.1) (FRAGMENT).
GN SOD.
OS Drosophila obscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97159564; PubMed=9007023;
RA Barrio E., Ayala F.J.;
RT "Evolution of the Drosophila obscura species group inferred from the
  Gpnh and Sod genes.";
RL Mol. Phylogenet. Evol. 7:79-93(1997).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
  CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) -> O(2) + H(2)O(2).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
DR EMBL; U47892; AAB50302.1;
DR HSSP; P00441; LSOS.
DR FlyBase; FBgn0016284; Dobs\Sod.
DR InterPro; IPR001424;
DR Pfam; PF00080; sodcu; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; PARTIAL.
KW Oxidoreductase; Copper; Zinc.
FT NON_TER 1 1
FT METAL 37 37 COPPER (BY SIMILARITY).
FT METAL 39 39 COPPER (BY SIMILARITY).
FT METAL 54 54 COPPER AND ZINC (BY SIMILARITY).
FT METAL 62 62 ZINC (BY SIMILARITY).
FT METAL 71 71 ZINC (BY SIMILARITY).
FT METAL 74 74 ZINC (BY SIMILARITY).
FT METAL 111 111 COPPER (BY SIMILARITY).
FT NON_TER 114 114
SQ SEQUENCE 114 AA; 11909 MW; 24F0EDF68C6A9D19 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 KVTGEV 285
Db 21 KVTGEV 26

RESULT 165
SODC_DROOB
ID SODC_DROOB STANDARD; PRT; 114 AA.
AC Q95035;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SUPEROXIDE DISMUTASE [CU-ZN] (EC 1.15.1.1) (FRAGMENT).
GN SOD.
OS Drosophila tolteca (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=30013;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97159564; PubMed=9007023;
RA Barrio E., Ayala F.J.;
RT "Evolution of the Drosophila obscura species group inferred from the
  Gpnh and Sod genes.";
RL Mol. Phylogenet. Evol. 7:79-93(1997).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
  CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) -> O(2) + H(2)O(2).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
DR EMBL; U47887; AAB50299.1;
DR HSSP; P15107; LXSQ.
DR FlyBase; FBgn0016233; Dmad\Sod.
DR InterPro; IPR001424;
DR Pfam; PF00080; sodcu; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; PARTIAL.
KW Oxidoreductase; Copper; Zinc.
FT NON_TER 1 1
FT METAL 37 37 COPPER (BY SIMILARITY).
FT METAL 39 39 COPPER (BY SIMILARITY).
FT METAL 54 54 COPPER AND ZINC (BY SIMILARITY).
FT METAL 62 62 ZINC (BY SIMILARITY).
FT METAL 71 71 ZINC (BY SIMILARITY).
FT METAL 74 74 ZINC (BY SIMILARITY).
FT METAL 111 111 COPPER (BY SIMILARITY).
FT NON_TER 114 114
SQ SEQUENCE 114 AA; 11934 MW; 24F0EDF68C6A9D19 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 KVTGEV 285
Db 21 KVTGEV 26

RESULT 166
SODC_DROTO
ID SODC_DROTO STANDARD; PRT; 114 AA.
AC Q95035;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SUPEROXIDE DISMUTASE [CU-ZN] (EC 1.15.1.1) (FRAGMENT).
GN SOD.
OS Drosophila tolteca (Fruit fly).

```

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7259;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NDSSC 14012-0201.0;  
RX MEDLINE=97159564; PubMed=9007023;  
RA Barrio E., Ayala F.J.;  
RT "Evolution of the Drosophila obscura species group inferred from the  
Gpdh and Sod genes.";  
RL Mol. Phylogenet. Evol. 7:79-93(1997).  
CC -!- FUNCTION: DESROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.  
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U47867; AAB50312.1; -  
DR HSSP; P15107; 1XSQ.  
DR FlyBase; FBgn0016390; Dtol\Sod.  
DR InterPro; IPR001424; -  
DR Pfam; PF00080; sodcu; 1.  
DR PROSITE; PS00087; SOD\_CU\_ZN\_1; FALSE\_NEG.  
DR PROSITE; PS00332; SOD\_CU\_ZN\_2; PARTIAL.  
KW Oxidoreductase; Copper; Zinc.  
FT NON\_TER 1 1  
FT METAL 37 37 COPPER (BY SIMILARITY).  
FT METAL 39 39 COPPER (BY SIMILARITY).  
FT METAL 54 54 COPPER AND ZINC (BY SIMILARITY).  
FT METAL 62 62 ZINC (BY SIMILARITY).  
FT METAL 71 71 ZINC (BY SIMILARITY).  
FT METAL 74 74 ZINC (BY SIMILARITY).  
FT METAL 111 111 COPPER (BY SIMILARITY).  
FT METAL 114 114  
FT NON\_TER 114 114  
SQ SEQUENCE 114 AA; 11763 MW; BAC26BC79DF2BBA3 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 114;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 280 KVTGEV 285  
DB 21 KVTGEV 26  
IIIIII  
  
RESULT 167  
VMEM\_PVX  
ID VMEM\_PVX STANDARD; PRT; 115 AA.  
AC P07697;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE 12 KDA MEMBRANE PROTEIN (ORF 3).  
OS Potato virus X (PVX).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.  
OX NCBI\_TaxID=12183;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89083520; PubMed=3205733;  
RA Skryabin K.G., Kraev A.S., Rozanov M.N., Chernov B.K.,  
RA Lukashova L.I., Atabekov J.G.;  
RT "The nucleotide sequence of potato virus X RNA.";

RL Nucleic Acids Res. 16:10929-10930(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Morozov S.Y., Lukashova L.I., Chernov B.K., Skryabin K.G.,  
RA Atabekov J.G.;  
RT "Nucleotide sequence of the open reading frames adjacent to the coat  
RT protein cistron in potato virus X genome.";  
RL FEBS Lett. 213:438-442(1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=X3;  
RX MEDLINE=88299944; PubMed=3404114;  
RA Huisman M.J., Linthorst H.J.M., Bol J.F., Cornelissen B.J.C.;  
RT "The complete nucleotide sequence of potato virus X and its  
RT homologues at the amino acid level with various plus-stranded RNA  
RT viruses.";  
RL J. Gen. Virol. 69:1789-1798(1988).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: TO ORF3 PROTEIN FROM OTHER POTEXVIRUSES AND TO 12 KDA  
CC PROTEIN FROM CARLAVIRUSES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M72416; AAA47169.1; -  
DR EMBL; X05198; CAA28828.1; -  
DR EMBL; D00344; BAA00251.1; -  
DR PIR; JAO104; WMWGP3.  
DR InterPro; IPR001896; -  
DR Pfam; PF01307; Plant\_vir\_prot; 1.  
KW Transmembrane.  
FT TRANSMEM 14 34 POTENTIAL.  
FT TRANSMEM 75 95 POTENTIAL.  
SQ SEQUENCE 115 AA; 12338 MW; 029983C61B842C0D CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 115;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 22 ALVSIT 27  
DB 26 ALVSIT 31  
IIIIII  
  
RESULT 168  
CH10\_MYCPN  
ID CH10\_MYCPN STANDARD; PRT; 116 AA.  
AC P75205;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE PUTATIVE 10 KDA CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).  
GN MOPB OR GROES OR MPN574 OR MP268.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
RA Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
RT pneumoniae";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -!- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES

CC THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).  
CC -!- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING  
CC (BY SIMILARITY)  
CC -!- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY. LOW SIMILARITY  
CC COMPARED TO OTHER BACTERIAL AND EUKARYOTIC CPN10.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AE000026; AAB95916.1; -  
CC InterPro: IPR001476; -  
CC Pfam: PF00166; CPN10; 2.  
CC PRINTS: PR00297; CHAPERONIN10.  
CC PROSITE: PS00681; CHAPERONINS\_CP10; FALSE\_NEG.  
CC Chapterone.  
CC KW  
CC SQ SEQUENCE 116 AA; 12618 MW; DBC4C83372334E52 CRC64;  
CC -----  
CC Query Match 0.5%; Score 6; DB 1; Length 116;  
CC Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC -----  
CC QY 347 DKNSA 352  
CC Db |||||  
CC 35 DKNSA 40  
CC -----  
CC RESULT 169  
CC REV\_HV1A2 STANDARD; PRT; 116 AA.  
CC AC P04623;  
CC DT 13-AUG-1987 (Rel. 05, Created)  
CC DT 13-AUG-1987 (Rel. 05, Last sequence update)  
CC DT 01-JUL-1993 (Rel. 26, Last annotation update)  
CC DE REV PROTEIN (ANTI-REPRESSION TRANSACTIVATOR PROTEIN) (ART/TRS).  
CC GN REV.  
CC OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).  
CC OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
CC OX NCBI\_TaxID=11685;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.  
CC RX MEDLINE=85090453; PubMed=2578227;  
CC RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,  
CC Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,  
CC Levy J.A., Dina D., Luciw P.A.;  
CC RT "Nucleotide sequence and expression of an AIDS-associated retrovirus  
CC (ARV-2).";  
CC RL Science 227:484-492(1985).  
CC CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE  
CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.  
CC CC -!- SUBCELLULAR LOCATION: NUCLEAR; ACCUMULATES IN THE NUCLEOLI.  
CC -!- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED  
CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: K02007; AAB59880.1; -  
CC HIV; K02007; REV3SF2.  
CC InterPro: IPR000625; -  
CC Pfam: PF00424; REV; 1.  
CC KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.  
CC SQ SEQUENCE 116 AA; 12966 MW; CD512C48A3B317A9 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
QY 908 NSGAQG 913  
Db |||||  
91 NSGAQG 96  
-----  
RESULT 170  
SMS\_CHICK STANDARD; PRT; 116 AA.  
ID SMS\_CHICK  
AC P33094;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE SOMATOSTATIN PRECURSOR [CONTAINS: SOMATOSTATIN-28; SOMATOSTATIN-14].  
GN SST.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RA Nata K., Kobayashi T., Karahashi K., Kato S., Yamamoto H.,  
RA Yonekura H., Okamoto H.;  
RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X60191; CAA42747.1; -  
CC FIR; S20630; S20630.  
CC KW Cleavage on pair of basic residues; Hormone; Signal.  
CC SIGNAL 1 24 BY SIMILARITY.  
CC FT PROPEP 25 88 BY SIMILARITY.  
CC FT PEPTIDE 89 116 SOMATOSTATIN-28.  
CC FT PEPTIDE 103 116 SOMATOSTATIN-14.  
CC FT DISULFID 105 116  
CC SQ SEQUENCE 116 AA; 12675 MW; 8A5BB9BDA8A291BA CRC64;  
-----  
Query Match 0.5%; Score 6; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
QY 53 AVGTVS 58  
Db |||||  
18 AVGTVS 23  
-----  
RESULT 171  
NLTD\_BRAOL STANDARD; PRT; 118 AA.  
ID NLTD\_BRAOL  
AC Q43304;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE NONSPECIFIC LIPID-TRANSFER PROTEIN D PRECURSOR (LTP D) (WAX-ASSOCIATED  
DE PROTEIN 9D).  
GN WAX9D.

OS Brassica oleracea (Cauliflower).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3712;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leaf;  
RX MEDLINE=94263227; PubMed=8203911;  
RA Pyee J., Yu H., Kolattukudy P.E.;  
RT "Identification of a lipid transfer protein as the major protein in  
RT the surface wax of broccoli (Brassica oleracea) leaves.";  
RL Arch. Biochem. Biophys. 311:460-468(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GREEN SPROUTING; TISSUE=Leaf;  
RX MEDLINE=95201828; PubMed=7894511;  
RA Pyee J., Kolattukudy P.E.;  
RT "The gene for the major cuticular wax-associated protein and three  
RT homologous genes from broccoli (Brassica oleracea) and their  
RT expression patterns";  
RL Plant J. 7:49-59(1995).  
CC -!- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER  
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY  
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING  
CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES.  
CC -!- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L29767; AAA32995.1; -  
DR EMBL; L33907; AAA73948.1; -  
DR HSP; P19656; IAPF.  
DR InterPro; IPR000528; -  
DR Pfam; PF00279; LTP; 1.  
DR PRINTS; PR00382; LIPIDTRNSFR.  
DR PROSITE; PS00597; PLANT\_LTP; 1.  
KW Lipid-binding; Transport; Signal; Multigene family.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 118 NONSPECIFIC LIPID-TRANSFER PROTEIN D.  
FT DISULFID 29 77 POTENTIAL.  
FT DISULFID 39 54 POTENTIAL.  
FT DISULFID 55 100 POTENTIAL.  
FT DISULFID 75 114 POTENTIAL.  
SQ SEQUENCE 118 AA; 11937 MW; 53214BCDC4491DFC CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 863 LAQNAP 868  
Db 43 LAQNAP 48  
|||||  
  
RESULT 172  
VPX\_SIVAT STANDARD; PRT; 119 AA.  
AC P05918;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-OCT-1989 (Rel. 12, Last annotation update)  
DE VPX PROTEIN (X ORF PROTEIN).  
GN VPX  
OS Simian immunodeficiency virus (TYO-1 isolate) (SIV-AGM).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11731;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88232906; PubMed=3374586;  
RA Fukasawa M., Miura T., Hasegawa A., Morikawa S., Tsujimoto H.,  
RA Miki K., Kitamura T., Hayami M.;  
RT "Sequence of simian immunodeficiency virus from African green monkey,  
RT a new member of the HIV/SIV group.";  
RL Nature 333:457-461(1988).  
CC -!- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X07805; CAA30660.1; -  
DR PIR; D30045; ASLJX4.  
DR HIV; X07805; VPXSGMTY.  
DR InterPro; IPR000012; -  
DR Pfam; PF00522; VPR; 1.  
DR PRINTS; PR00444; HIVPRVPX.  
KW AIDS.  
SQ SEQUENCE 119 AA; 14340 MW; 27D992D352AEC276 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 119;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1129 LQDLNQ 1134  
Db 33 LQDLNQ 38  
|||||  
  
RESULT 173  
YHHM\_ECOLI STANDARD; PRT; 119 AA.  
AC P37615;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL 13.5 KDA PROTEIN IN FTSY-NIKA INTERGENIC REGION (FI19).  
GN YHHM.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=94316500; PubMed=8041620;  
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;  
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the  
RT region from 76.0 to 81.5 minutes.";  
RL Nucleic Acids Res. 22:2576-2586(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U00039; AAB18442.1; -  
DR EMBL; AE000422; AAC76492.1; -  
DR EcoGene; EG12213; yhhm.  
KW Hypothetical protein.  
SQ SEQUENCE 119 AA; 13496 MW; CB4A67C204FG33E9 CRC64;

```

Query Match          0.5%; Score 6; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 LTVGDK 295
      |||||
DB 97 LTVGDK 102

RESULT 174
YRF4_SHIFL
ID YRF4_SHIFL STANDARD; PRT; 119 AA.
AC P37790;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE HYPOTHETICAL 13.8 KDA PROTEIN IN RFBJ-GND INTERGENIC REGION (ORF12X8).
OS Shigella flexneri.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Shigella.
CC NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE 2A;
RX MEDLINE=94131953; PubMed=7507920;
RA Morona R., Mavris M., Fallarino A., Manning P.A.;
RT "Characterization of the rfc region of Shigella flexneri.";
RL J. Bacteriol. 176:733-747(1994).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X71970; CAA50780.1; -
KW Hypothetical protein.
SQ SEQUENCE 119 AA; 13766 MW; 1337D6E61DAD145D CRC64;

Query Match          0.5%; Score 6; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 GNAIGD 796
      |||||
DB 44 GNAIGD 49

RESULT 175
RL24_ARCFU
ID RL24_ARCFU STANDARD; PRT; 120 AA.
AC O28365;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L24P.
GN RPL24P OR AF1914.
OS Archaeoglobus fulgidus.
CC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
CC Archaeoglobus.
CC NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervlavage A.R., Graham D.E., Kyrpides N.C.,

```

```

RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC
CC -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000971; AAB89339.1; -
DR TIGR; AF1914; -
DR InterPro; IPR000302; -
DR Pfam; PF00467; Ribosomal_L24; 1.
DR PROSITE; PS01108; RIBOSOMAL_L24; 1.
KW Hypothetical protein; Ribosomal protein.
SQ SEQUENCE 120 AA; 13924 MW; 1F754A70D5ABCFL9 CRC64;

Query Match          0.5%; Score 6; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 LEVDMK 137
      |||||
DB 69 LEVDMK 74

RESULT 176
YGC9_YEAST
ID YGC9_YEAST STANDARD; PRT; 120 AA.
AC P53188;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL 14.4 KDA PROTEIN IN RPL30-CWH41 INTERGENIC REGION.
GN YGL029W.
OS Saccharomyces cerevisiae (Baker's yeast).
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Hebling U., Hofmann B., Delius H.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z72551; CAA96730.1; -
DR HSP; P09012; 1FHT.
DR SGD; S0002997; CGR1.
KW Hypothetical protein.
SQ SEQUENCE 120 AA; 14417 MW; E2750E786489E9CA CRC64;

Query Match          0.5%; Score 6; DB 1; Length 120;

```

Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 762 QFKERL 767  
Db 56 QFKERL 61

## RESULT 177

VIB2\_AGR5  
ID VIB2\_AGR5 STANDARD; PRT; 121 AA.  
AC P17792;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-NOV-1990 (Rel. 16, Last annotation update)  
DE VIB2 PROTEIN PRECURSOR.  
OS VIB2.  
GN Agrobacterium tumefaciens.  
OG Plasmid pTIC58  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Agrobacterium.  
OX NCBI\_TaxID=362;  
RN [1]  
RP MEDLINE=90318324; PubMed=2370849;  
RX Kuldau G.A., de Vos G., Owen J., McCaffrey G., Zambryski P.;  
RA "The virB operon of Agrobacterium tumefaciens pTIC58 encodes 11 open  
RT reading frames.";  
RL Mol. Gen. Genet. 221:256-266(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90301800; PubMed=2194232;  
RA Rogovsky P.M., Powell B.S., Shirasu K., Lin T.-S., Morel P.,  
RA Zyprian E.M., Steck T.R., Kado C.I.;  
RT "Molecular characterization of the vir regulon of Agrobacterium  
tumefaciens: complete nucleotide sequence and gene organization of  
the 28.63-kbp regulon cloned as a single unit.";  
RL Plasmid 23:85-106(1990).  
CC -!- FUNCTION: VIRB PROTEINS ARE SUGGESTED TO ACT AT THE BACTERIAL  
CC SURFACE AND THERE PLAY AN IMPORTANT ROLE IN DIRECTING T-DNA  
CC TRANSFER TO PLANT CELLS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X53264; CAA37355.1; -;  
DR EMBL; J03320; AAA91592.1; -;  
DR PIR; S12342; B2AG58.  
KW Crown gall tumor; Plasmid; Signal.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 121 VIRB2 PROTEIN.  
SQ SEQUENCE 121 AA; 12318 MW; 0499B92984A417AE CRC64;

Query Match 0.5%; Score 6; DB 1; Length 121;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 LSLSNA 974  
Db 14 LSLSNA 19

## RESULT 178

YEDR\_ECOLI  
ID YEDR\_ECOLI STANDARD; PRT; 121 AA.  
AC P76334;  
DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL 13.8 KDA PROTEIN IN DCM-SHIA INTERGENIC REGION.  
GN YEDR.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90301800; PubMed=2194232;  
RA Rogovsky P.M., Powell B.S., Shirasu K., Lin T.-S., Morel P.,  
RA Zyprian E.M., Steck T.R., Kado C.I.;  
RT "Molecular characterization of the vir regulon of Agrobacterium  
tumefaciens: complete nucleotide sequence and gene organization of  
the 28.63-kbp regulon cloned as a single unit.";  
RL Plasmid 23:85-106(1990).  
CC -!- FUNCTION: VIRB PROTEINS ARE SUGGESTED TO ACT AT THE BACTERIAL  
CC SURFACE AND THERE PLAY AN IMPORTANT ROLE IN DIRECTING T-DNA  
CC TRANSFER TO PLANT CELLS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AE000288; AAC75029.1; ALT\_INIT.  
DR EcoGene; EG14041; Yedr.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 38 58 POTENTIAL.  
FT TRANSMEM 86 106 POTENTIAL.  
SQ SEQUENCE 121 AA; 13838 MW; 30619D75B31A1E70 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 121;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ALVGAL 23  
Db 97 ALVGAL 102

RESULT 179  
MERR\_THIFE  
ID MERR\_THIFE STANDARD; PRT; 122 AA.  
AC P22896;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 01-OCT-1993 (Rel. 27, Last annotation update)  
DE MERCURIC RESISTANCE OPERON REGULATORY PROTEIN (FRAGMENT).  
GN MERR.  
OS Thiobacillus ferrooxidans.  
OG Plasmid pTF-FC2.  
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.  
OX NCBI\_TaxID=920;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93015664; PubMed=1400173;  
RA Rohrer J., Rawlings D.E.;  
RT "Sequence analysis and characterization of the mobilization region of  
RT a broad-host-range plasmid, pTF-FC2, isolated from Thiobacillus  
RT ferrooxidans.";  
RL J. Bacteriol. 174:6230-6237(1992).  
CC -!- FUNCTION: MEDIATES THE MERCURIC-DEPENDENT INDUCTION OF MERCURY  
CC RESISTANCE OPERON. IN THE ABSENCE OF MERCURY MERR REPRESSES  
CC TRANSCRIPTION BY BINDING TIGHTLY TO THE MERR OPERATOR REGION;  
CC WHEN MERCURY IS PRESENT THE DIMERIC COMPLEX BINDS A SINGLE ION  
CC AND BECOMES A POTENT TRANSCRIPTIONAL ACTIVATOR, WHILE REMAINING  
CC BOUND TO THE MERR SITE.  
CC -!- SIMILARITY: BELONGS TO THE MERR FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M57717; AAA27396.1; -  
DR PIR; I43256; I43256.  
DR PIR; S27629; S27629.  
DR InterPro: IPR000551; -  
DR Pfam; PF00376; merR; 1.  
DR PROSITE; PS00552; HTH\_MERR\_FAMILY; 1.  
KW Transcription regulation; Activator; Repressor; Mercuroic resistance;  
KW Mercury; DNA-binding; Plasmid.  
FT DNA\_BIND 8 27  
FT METAL 81 81 HG(2+) (BY SIMILARITY) (OR 82).  
FT METAL 117 117 HG(2+) (BY SIMILARITY).  
FT NON\_TER 122 122  
SQ SEQUENCE 122 AA; 13617 MW; D067FD046497CBE3 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 122;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 690 GKVATL 695  
DB 9 GKVATL 14  
|||||  
-----  
RESULT 180  
NB8M\_CABEL STANDARD; PRT; 123 AA.  
AC P90789;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE PROBABLE NADH-UBIQUINONE OXIDOREDUCTASE B18 SUBUNIT (EC 1.6.5.3)  
DE (EC 1.6.99.3) (COMPLEX I-B18) (CI-B18).  
GN D2030.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Wilkinson J.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED  
CC TO BE UBIQUINONE (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS (BY  
CC SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; 273906; CAA98116.1; -  
DR WormPep; D2030.4; CE09081.  
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
SQ SEQUENCE 123 AA; 14408 MW; 1DA8858527116F09 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 123;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 865 QNAPFA 870  
DB 75 QNAPFA 80  
|||||  
-----  
RESULT 181  
WN7A\_ALOVU STANDARD; PRT; 123 AA.  
AC P28105;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE WNT-7A PROTEIN (FRAGMENT).  
GN WNT-7A.  
OS Alopias vulpinus (Thresher shark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Lamniformes; Alopiidae;  
OC Alopiinae; Alopias.  
OX NCBI\_TaxID=7852;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92279273; PubMed=1534411;  
RA Sidow A.;  
RT "Diversification of the Wnt gene family on the ancestral lineage of  
RT vertebrates".  
RL Proc. Natl. Acad. Sci. U.S.A. 89:5098-5102(1992).  
CC -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING  
CC MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF  
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.  
CC -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE  
CC EXTRACELLULAR MATRIX.  
CC -!- SIMILARITY: BELONGS TO THE WNT FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M91256; AAA48542.1; -  
DR InterPro: IPR000970; -  
DR Pfam; PF00110; wnt; 1.  
DR PROSITE; PS00246; WNT1; PARTIAL.  
KW Developmental protein; Glycoprotein.  
FT NON\_TER 1 1  
FT NON\_TER 123 123  
SQ SEQUENCE 123 AA; 14274 MW; A14C64948BBB1DA4 CRC64;  
  
Query Match 0.5%; Score 6; DB 1; Length 123;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 359 SSQNS 364  
DB 93 SSQNS 98  
|||||  
-----  
RESULT 182  
RL17\_MYCPN STANDARD; PRT; 124 AA.  
AC Q59547;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE 50S RIBOSOMAL PROTEIN L17.

GN RPLQ OR MPN192 OR MP639.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE=96177562; PubMed=8604303;  
 RA Hilbert H., Himmelreich R., Plagens H., Herrmann R.;  
 RT "Sequence analysis of 56 kb from the genome of the bacterium  
 RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and  
 RT a cluster of ribosomal protein genes.";  
 RL Nucleic Acids Res. 24:628-639(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -1- SIMILARITY: BELONGS TO THE L17P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U34795; AAC3689.1; -  
 DR EMBL; AE000061; AAB96287.1; -  
 DR InterPro; IPR000456; -  
 DR Pfam; PF01196; Ribosomal\_L17; 1.  
 DR PROSITE; PS01167; RIBOSOMAL\_L17; 1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 124 AA; 14245 MW; 3A627DB7EBF8C62E CRC64;  
 Query Match 0.5%; Score 6; DB 1; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1021 APKYK 1026  
 DB 87 APKYK 92  
 RESULT 183  
 ID ACTS\_PLEWA STANDARD; PRT; 125 AA.  
 AC P10994;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ACTIN, ALPHA SKELETAL MUSCLE (FRAGMENT).  
 OS Pleurodeles waltlii (Iberian ribbed newt).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;  
 OC Pleurodeles.  
 OX NCBI\_TaxID=8319;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=87254218; PubMed=3453122;  
 RX Khrestchatisky M., Fontes M.;  
 RA "There is an alpha-actin skeletal muscle-specific gene in a  
 RT salamander (Pleurodeles waltlii).";  
 RL J. Mol. Biol. 193:409-412(1987).  
 CC -1- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED  
 CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED

CC IN ALL EUKARYOTIC CELLS.  
 CC -1- SUBUNIT: POLYMERIZATION OF GLOBULAR ACTIN (G-ACTIN) LEADS TO A  
 CC STRUCTURAL FILAMENT (F-ACTIN) IN THE FORM OF A TWO-STRANDED  
 CC HELIX. EACH ACTIN CAN BIND TO 4 OTHERS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X05106; CAA28753.1; -  
 DR PIR; A26337; A26337.  
 DR HSP; P02568; IATN.  
 DR InterPro; IPR000279; -  
 DR Pfam; PF00022; actin; 1.  
 DR PROSITE; PS00406; ACTINS\_1; PARTIAL.  
 DR PROSITE; PS00432; ACTINS\_2; 1.  
 DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; PARTIAL.  
 KW Multigene family; Structural protein; Muscle protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 125 AA; 14102 MW; 4012AD024CD2AC89 CRC64;  
 Query Match 0.5%; Score 6; DB 1; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 323 IAPPE 328  
 DB 79 IAPPE 84  
 RESULT 184  
 ID C556\_AGRTA STANDARD; PRT; 125 AA.  
 AC P00141;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CYTOCHROME C-556 (C556).  
 OS Agrobacterium tumefaciens (strain apple 185).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Agrobacterium.  
 OX NCBI\_TaxID=370;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=83287429; PubMed=6309523;  
 RA Tempst P., van Beeumen J.;  
 RT "The amino acid sequence of cytochrome c-556 from Agrobacterium  
 RT tumefaciens strain Apple 185.";  
 RL Eur. J. Biochem. 135:321-330(1983).  
 CC -1- FUNCTION: CYTOCHROME C556 IS A LOW-SPIN, MONOMERIC MONOHEME  
 CC PROTEIN.  
 CC PIR; A00134; CCAGA6.  
 DR HSP; P00147; ICPR.  
 DR InterPro; IPR000345; -  
 DR InterPro; IPR002321; -  
 DR Pfam; PF01322; Cytochrome\_C\_2; 1.  
 DR PRINTS; PR00608; CYTOCHROME\_CII.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 KW Electron transport; Heme.  
 FT METAL 13  
 FT BINDING 113 113 IRON (HEME AXIAL LIGAND).  
 FT BINDING 116 116 HEME (COVALENT).  
 FT METAL 117 117 IRON (HEME AXIAL LIGAND).  
 SQ SEQUENCE 125 AA; 12985 MW; 1E3240764EE7A508 CRC64;



```
Query Match          0.5%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 KASLT 422
DB 39 KASLT 44

RESULT 185
RPA9_YEAST          STANDARD; PRT; 125 AA.
AC P32529;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-DIRECTED RNA POLYMERASE I 13.7 KDA POLYPEPTIDE (EC 2.7.7.6)
DE (A12.2).
GN RPA12 OR RRN4 OR YJR063W OR J1747.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93109294; PubMed=8417319;
RA Nogai Y., Yano R., Dodd J., Carles C., Nomura M.;
RT "Gene RRN4 in Saccharomyces cerevisiae encodes the A12.2 subunit of
RT RNA polymerase I and is essential only at high temperatures."
RL Mol. Cell. Biol. 13:114-122(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96437976; PubMed=8840504;
RA Huang M.-E., Manus V., Chuat J.-C., Galibert F.;
RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
RT reading frames and a gene cluster with a counterpart on chromosome
RT XI."
RL Yeast 12:869-875(1996).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
CC RNA(N).
CC -!- SUBUNIT: RNA POLYMERASE I CONSISTS OF 14 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND 5.8S rRNA GENES.
CC -!- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOM / EUKARYOTIC RPA12/
CC RPB9 / RPB11 RNA POLYMERASE FAMILY.
CC -!- SIMILARITY: BELONGS TO THE TRANSCRIPTION FACTOR S-II FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L00708; AAA34992.1; -
CC DR EMBL; L35564; AAB59319.1; -
CC DR EMBL; Z49563; CAA89591.1; -
CC DR EMBL; L47993; AAB39289.1; -
CC DR PIR; A48107; A48107.
CC DR PIR; S47937; S47937.
CC DR HSP; Q56254; IQYP.
CC DR SGD; S0003824; RPA12.
CC DR InterPro; IPR001222; -
CC DR InterPro; IPR001529; -
CC DR Pfam; PF01096; TFIIS; 1.

DR PROSITE; PS00466; TFIIS; 1.
DR PROSITE; PS01030; RNA_POL15KD; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription;
KW Nuclear protein; Zinc-finger. C4-TYPE (POTENTIAL).
FT 2N_FING 10 33
FT 2N_FING 86 117
FT ZN_RIBBON (POTENTIAL).
SQ SEQUENCE 125 AA; 13660 MW; D79372070819987C CRC64;

Query Match          0.5%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 665 S0FSNL 670
DB 40 S0FSNL 45

RESULT 186
SMD3_HUMAN          STANDARD; PRT; 126 AA.
ID SMD3_HUMAN
AC P43331;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3 (SNRNP CORE PROTEIN D3) (SM-D3).
GN SNRPD3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=95083692; PubMed=7527560;
RA Lehmeier T., Raker V., Hermann H., Luehrmann R.;
RT "cDNA cloning of the Sm proteins D2 and D3 from human small nuclear
RT ribonucleoproteins: evidence for a direct D1-D2 interaction."
RL Proc. Natl. Acad. Sci. U.S.A. 91:12317-12321(1994).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- DISEASE: IN THE AUTOIMMUNE DISEASE SYSTEMIC LUPUS ERYTHEMATOSUS,
CC ANTINUCLEAR ANTIBODIES ARE DEVELOPED WITH SM SPECIFICITY.
CC -!- SIMILARITY: BELONGS TO THE SNRNP CORE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U15009; AAA57034.1; -
CC DR EMBL; 601062; -
CC DR InterPro; IPR001163; -
CC DR Pfam; PF01423; Sm; 1.
CC DR Systemic lupus erythematosus; Repeat. mRNA splicing; mRNA processing;
CC Systemic lupus erythematosus; Repeat. ARG/LYS-RICH (BASIC).
FT DOMAIN 84 126
FT DOMAIN 110 119
FT DOMAIN 5 X 2 AA TANDEM REPEATS OF [RM]-G.
SQ SEQUENCE 126 AA; 13916 MW; 59A6E78E60AA3E0 CRC64;

Query Match          0.5%; Score 6; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 GSGAGR 213
DB 92 GSGAGR 97

RESULT 187
YMF1_ECOLI
```

ID YMF1\_ECOLI STANDARD; PRT; 128 AA.  
 AC P75972;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOTHETICAL 14.5 KDA PROTEIN IN INTE-PIN INTERGENIC REGION.  
 GN YMF1.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=362;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474 (1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF000214; AAC74227.1;  
 DR EcoGene; EGI4247; ymf1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 128 AA; 14513 MW; 2E5216E84BCEBBFB CRC64;  
  
 Query Match 0.5%; Score 6; DB 1; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1052 SAGVDA 1057  
 Db 85 SAGVDA 90  
 |||||  
  
 RESULT 188  
 LYC2\_CANFA STANDARD; PRT; 130 AA.  
 ID LYC2\_CANFA  
 AC P81709;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE LYSOZYME C, SPLEEN ISOZYME (EC 3.2.1.17) (1,4-BETA-N-ACETYLURAMIDASE  
 DE C).  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Spleen;  
 RX MEDLINE=94361523; PubMed=8080284;  
 RA Grobler J.A., Rao K.R., Pervais S., Brew K.;  
 RT "Sequences of two highly divergent canine type c lysozymes:  
 RT implications for the evolutionary origins of the lysozyme/alpha-  
 RT lactalbumin superfamily.";  
 RL Arch. Biochem. Biophys. 313:360-366(1994).  
 CC -1- FUNCTION: LYSOZYMES HAVE PRIMARILY BACTERIOLYTIC FUNCTION; THOSE  
 CC IN TISSUES AND BODY FLUIDS ARE ASSOCIATED WITH THE MONOCYTE-  
 CC MACROPHAGE SYSTEM AND ENHANCE THE ACTIVITY OF IMMUNOAGENTS (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES BETWEEN  
 CC N-ACETYL-D-GLUCOSAMINE AND N-ACETYLURAMIC ACID IN PEPTIDOLYCAN

CC HETEROPOLYMERS OF THE PROKARYOTES CELL WALLS.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: LYSOZYME C IS CAPABLE OF BOTH HYDROLYSIS AND  
 CC TRANSGLYCOSYLATION; IT SHOWS ALSO A SLIGHT ESTERASE ACTIVITY. IT  
 CC ACTS RAPIDLY ON BOTH PEPTIDE-SUBSTITUTED AND UNSUBSTITUTED  
 CC PEPTIDOLYCAN S, SLOWLY, ON CHITIN OLIGOSACCHARIDES (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 22 OF GLYCOSYL HYDROLASES.  
 DR Pfam; PF00062; lys; 1.  
 DR PROSITE; PS00128; LACTALBUMIN\_LYSOZYME; 1.  
 KW Hydrolase; Glycosidase; Bacteriolytic enzyme; Multigene family.  
 FT DISULFID 6 128 BY SIMILARITY.  
 FT DISULFID 30 116 BY SIMILARITY.  
 FT DISULFID 65 81 BY SIMILARITY.  
 FT ACT\_SITE 77 95 BY SIMILARITY.  
 FT ACT\_SITE 35 35 BY SIMILARITY.  
 FT ACT\_SITE 53 53 BY SIMILARITY.  
 SQ SEQUENCE 130 AA; 14578 MW; 96C9BA30478D60F6 CRC64;  
  
 Query Match 0.5%; Score 6; DB 1; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1126 SALLQD 1131  
 Db 82 SALLQD 87  
 |||||  
  
 RESULT 189  
 LYCK\_SHEEP STANDARD; PRT; 130 AA.  
 ID LYCK\_SHEEP  
 AC P80190;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE LYSOZYME C, KIDNEY (EC 3.2.1.17) (1,4-BETA-N-ACETYLURAMIDASE C).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Kidney;  
 RX MEDLINE=93238751; PubMed=8477739;  
 RA Ito Y., Yamada H., Nakamura M., Yoshikawa A., Ueda T., Imoto T.;  
 RT "The primary structures and properties of non-stomach lysozymes of  
 RT sheep and cow, and implication for functional divergence of  
 RT lysozyme.";  
 RL Eur. J. Biochem. 213:649-658(1993).  
 RN [2]  
 RP SEQUENCE OF 34-63 FROM N.A.  
 RX MEDLINE=96054033; PubMed=7563116;  
 RA Irwin D.M.;  
 RT "Evolution of the bovine lysozyme gene family: changes in gene  
 RT expression and reversion of function.";  
 RL J. Mol. Evol. 41:299-312(1995).  
 CC -1- FUNCTION: LYSOZYMES HAVE PRIMARILY BACTERIOLYTIC FUNCTION; THOSE  
 CC IN TISSUES AND BODY FLUIDS ARE ASSOCIATED WITH THE MONOCYTE-  
 CC MACROPHAGE SYSTEM AND ENHANCE THE ACTIVITY OF IMMUNOAGENTS.  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES BETWEEN  
 CC N-ACETYL-D-GLUCOSAMINE AND N-ACETYLURAMIC ACID IN PEPTIDOLYCAN  
 CC HETEROPOLYMERS OF THE PROKARYOTES CELL WALLS.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN HERE IS A NON-STOMACH ISOZYME.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 22 OF GLYCOSYL HYDROLASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>;  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U19473; AAB85547.1; -  
 DR PIR: S32465; S32465.  
 DR HSP: P00695; I0UF.  
 DR InterPro; IPR000974; -  
 DR InterPro; IPR001916; -  
 DR Pfam: PF00062; Lys: 1.  
 DR PRINTS: PR00135; LY2LACT.  
 DR PRINTS: PR00137; LYSOZYME.  
 DR PROSITE; PS00128; LACTALBUMIN LYSOZYME; 1.  
 KW Hydrolase; Glycosidase; Bacteriolytic enzyme; Multigene family.  
 FT DISULFID 6 128 BY SIMILARITY.  
 FT DISULFID 30 116 BY SIMILARITY.  
 FT DISULFID 65 81 BY SIMILARITY.  
 FT DISULFID 77 95 BY SIMILARITY.  
 FT ACT\_SITE 35 35 BY SIMILARITY.  
 FT ACT\_SITE 53 53 BY SIMILARITY.  
 SQ SEQUENCE 130 AA; 14611 MW; C79D70D3B9F70A8E CRC64;

Query Match 0.5%; Score 6; DB 1; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1126 SALLQD 1131  
 DB 82 SALLQD 87

RESULT 190  
 RS11\_METTH STANDARD; PRT; 130 AA.  
 AC O26143;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 30S RIBOSOMAL PROTEIN SLIP.  
 GN RPSLIP OR MTH36.

OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanobacter.  
 OX NCBI\_TaxID=145262;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-DELTA H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Alredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT delah: functional analysis and comparative genomics."  
 RL J. Bacteriol. 179:7135-7155(1997).

CC -!- SIMILARITY: BELONGS TO THE SLIP FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>;  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE000796; AAB84544.1; -  
 DR InterPro; IPR001971; -  
 DR Pfam: PF00411; Ribosomal\_S11; 1.  
 DR PROSITE; PS00054; RIBOSOMAL\_S11;  
 KW Ribosomal protein.

SQ SEQUENCE 130 AA; 13672 MW; DL1A46D458556A9 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1075 YSSFNN 1080  
 DB 13 YSSFNN 18

RESULT 191  
 RS11\_AERPE STANDARD; PRT; 131 AA.  
 AC O9YB55;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 30S RIBOSOMAL PROTEIN SLIP.  
 GN RPSLIP OR APE1742.

OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;  
 OC Aeropyrum.

OX NCBI\_TaxID=56636;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,  
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 RT crenarchaeon, Aeropyrum pernix K1."  
 RL DNA Res. 6:83-101(1999).

CC -!- SIMILARITY: BELONGS TO THE SLIP FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>;  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AP000062; BAA80743.1; -  
 DR InterPro; IPR001971; -  
 DR Pfam: PF00411; Ribosomal\_S11; 1.  
 DR PROSITE; PS00054; RIBOSOMAL\_S11;  
 KW Ribosomal protein.  
 SQ SEQUENCE 131 AA; 14021 MW; 9138AF1135AE8E35 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1075 YSSFNN 1080  
 DB 14 YSSFNN 19

RESULT 192  
 RS8\_ARCFU STANDARD; PRT; 131 AA.  
 ID RS8\_ARCFU  
 AC O28369;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE 30S RIBOSOMAL PROTEIN S8F.

```

GN RPS8P OR AF1910.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Fleischmann R.D., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: BELONGS TO THE S8P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000971; AB89341.1; -
DR TIGR; AF1910; -
DR InterPro; IPR000630; -
DR Pfam; PF00410; Ribosomal_S8; 1.
DR PROSITE; PS00053; RIBOSOMAL_S8; 1.
KW Ribosomal protein.
SQ SEQUENCE 131 AA; 14684 MW; E1B413B76449B3E CRC64;

Query Match 0.5%; Score 6; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 971 LSNAMI 976
DB 8 LSNAMI 13
|||||

RESULT 193
SYR_BUCAP STANDARD; PRT; 131 AA.
AC Q08888;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ARGINYL-TRNA SYNTHETASE (EC 6.1.1.19) (ARGININE--TRNA LIGASE) (ARGRS)
DE (FRAGMENT).
GN ARGS.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RC SEQUENCE FROM N.A.
RC MEDLINE=94131280; PubMed=7507875;
RA Munson M.A., Baumann L., Baumann P.;
RT "Buchnera aphidicola (a prokaryotic endosymbiont of aphids) contains
RT a putative 16S rRNA operon unlinked to the 23S rRNA-encoding gene;
RT sequence determination, and promoter and terminator analysis."
RL Gene 137:171-178(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-ARGININE + TRNA(ARG) -> AMP +

```

---

```

CC PYROPHOSPHATE + L-ARGINYL-TRNA(ARG).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L18927; AAA17432.1; -
DR InterPro; IPR001412; -
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; PARTIAL.
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 123 >131
FT NON_TER 131 131
SQ SEQUENCE 131 AA; 15596 MW; BA82EF89049BD586 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 719 LIKINS 724
DB 15 LIKINS 20
|||||

RESULT 194
SVP2_CAVPO STANDARD; PRT; 132 AA.
AC P16531;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE SEMINAL VESICLE PROTEIN SVP-2 PRECURSOR (FRAGMENT).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RC SEQUENCE FROM N.A., AND SEQUENCE OF 15-39.
RX MEDLINE=90114184; PubMed=2691882;
RA Hagstrom J.E., Harvey S., Madden B., McCormick D., Wieben E.D.;
RT "Androgens affect the processing of secretory protein precursors in
RT the guinea pig seminal vesicle. II. Identification of conserved sites
RT for protein processing."
RL Mol. Endocrinol. 3:1797-1806(1989).
CC -1- ALTERNATIVE PRODUCTS: AN SVP-2 THAT CONTAINS ALA AT POSITION 18
CC COULD BE THE PRODUCT OF A SECOND ALLELE OR IT COULD RESULT FROM
CC ALTERNATIVE RNA PROCESSING OF A TRANSCRIPT FROM A SINGLE GENE.
CC -1- MISCELLANEOUS: SVP-2 IS ONE OF THE 4 MAJOR SECRETORY PROTEINS,
CC SECRETED BY THE GUINEA PIG SEMINAL VESICLE EPITHELIUM (SVE).
CC -1- SIMILARITY: TO THE SVP-1/-3/-4 PRECURSOR, PARTICULARLY IN REGIONS
CC WHERE PROTEIN PROCESSING MUST OCCUR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31653; AAA37055.1; -
DR PIR; A41405; A41405.
KW Signal; Alternative splicing; Cleavage on pair of basic residues;
KW Seminal vesicle.
FT NON_TER 1 1
FT SIGNAL <1 14

```

```

FT CHAIN 15 103 SEMINAL VESICLE PROTEIN SVP-2.
FT PROPEP 104 132 POTENTIAL.
FT VARIANT 18 18 R -> RA.
FT VARIANT 95 95 S -> L.
SQ SEQUENCE 132 AA; 14926 MW; 714EEE7B8EE9CA95 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6e+02; Length 132;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 NOASGR 443
DB 10 NOASGR 15

RESULT 195
YJW_ECOLI
ID YJW_ECOLI STANDARD; PRT; 132 AA.
AC P39394;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 14.6 KDA PROTEIN IN MCRB-HSDS INTERGENIC REGION (F132).
GN YJW.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U14003; AAA97244.1; -.
CC DR EMBL; AE000505; AAC77303.1; -.
CC DR EcoGene; EG12584; yjiw.
KW Hypothetical protein.
SQ SEQUENCE 132 AA; 14576 MW; 3E53097CD17B0C62 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6e+02; Length 132;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 ATGTAV 54
DB 72 ATGTAV 77

RESULT 196
OTOR_RANCA
ID OTOR_RANCA STANDARD; PRT; 133 AA.
AC Q918P5;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OTORAPLIN PRECURSOR.
GN OTOR.
OS Rana catesbeiana (Bull frog).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=10873378;
RA Robertson N.G., Heller S., Lin J.S., Resendes B.L., Weremowicz S.,
RA Denis C.S., Bell A.M., Hudspeth A.J., Morton C.C.;
RT "A novel conserved cochlear gene, OTOR: identification, expression
RT analysis, and chromosomal mapping.";
RL Genomics 66:242-248(2000).
CC -!- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE MIA / OTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AF233519; AAF82728.1; -.
CC DR InterPro; IPR001452; -.
CC DR PROSITE; PS50002; SH3; 1.
KW SIGNAL; SH3 domain. 23
FT CHAIN 1 23 POTENTIAL.
FT CHAIN 24 133 OTORAPLIN.
FT DOMAIN 48 115 SH3.
FT DOMAIN 133 AA; 15243 MW; 25440C1A3CF911AE CRC64;
SQ SEQUENCE 133 AA; 15243 MW; 25440C1A3CF911AE CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6e+02; Length 133;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 501 RFVNLK 506
DB 59 RFVNLK 64

RESULT 197
LYGA_MOUSE
ID LYGA_MOUSE STANDARD; PRT; 134 AA.
AC P05533;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LYMPHOCYTE ANTIGEN LY-6A.2/LY-6E.1 PRECURSOR (T-CELL-ACTIVATING
DE PROTEIN) (TAP).
GN LY6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88031496; PubMed=2822573;
RA Palfree R.G.E., Leclair K.P., Bothwell A.L.M., Hammerling U.;
RT "cDNA characterization of an Ly-6.2 gene expressed in BW5147 tumor
RT cells.";
RL Immunogenetics 26:389-391(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88176923; PubMed=2895473;
RA Reiser H., Colligan J., Benacerraf B., Rock K.L.;
RT "Cloning and expression of a cDNA for the T-cell-activating protein
RT TAP.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:2255-2259(1988).
RN [3]
RP SEQUENCE FROM N.A. (LY-6E.1).
RX MEDLINE=67133482; PubMed=3028776;

```

RA Leclair K.P., Palfrey R.G.E., Flood P.M., Hammerling U.,  
RA Bothwell A.L.M.;  
RT "Isolation of a murine Ly-6 cDNA reveals a new multigene family.";  
RL EMBO J. 5:3227-3234(1986).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RX MEDLINE=90377204; PubMed=1697928;  
RA Khan K., Dad Lindwall G., Maher S.E., Bothwell A.L.M.;  
RT "Characterization of promoter elements of an interferon-inducible Ly-  
RT 6E/A differentiation antigen, which is expressed on activated T cells  
RT and hematopoietic stem cells.";  
RL Mol. Cell. Biol. 10:5150-5159(1990).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92250126; PubMed=1315719;  
RA Stanford W.L., Bruyns E., Snodgrass H.R.;  
RT "The isolation and sequence of the chromosomal gene and regulatory  
RT regions of Ly-6A.2.";  
RL Immunogenetics 35:408-411(1992).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91225510; PubMed=1709198;  
RA McGrew J.T., Rock K.L.;  
RT "Isolation, expression, and sequence of the TAP/Ly-6A.2 chromosomal  
RT gene.";  
RL J. Immunol. 146:3633-3638(1991).  
CC -!- FUNCTION: T-CELL ACTIVATION.  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
CC -!- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M18184; AAA39463.1; -  
DR EMBL; J03636; AAA40163.1; -  
DR EMBL; X04653; CAA28351.1; -  
DR EMBL; M37707; AAA39487.1; -  
DR EMBL; M74013; AAA39484.1; -  
DR EMBL; M73552; AAA39465.1; -  
DR EMBL; M59713; AAA40164.1; -  
DR PIR; A31935; A31935.  
DR PIR; A25708; A25708.  
DR PIR; A32506; A32506.  
DR PIR; A35921; A35921.  
DR HSP; P01379; ILSI.  
DR MGD; MGI:96880; LY6.  
DR InterPro: IPR001526; -  
DR Pfam: PF00021; UPAR\_LY6; 1.  
DR PROSITE; PS00983; LY6\_UPAR; 1.  
KW T-cell; signal; Antigen; Multigene family; GPI-anchor.  
FT SIGNAL 1 26  
FT CHAIN 27 2119 LYMPHOCYTE ANTIGEN LY-6A.2/LY-6E.1.  
FT PROPEP 2120 134 REMOVED IN NATURE FORM (POTENTIAL).  
FT DOMAIN 27 119 UPAR/LY6.  
FT DISULFID 29 53 BY SIMILARITY.  
FT DISULFID 32 41 BY SIMILARITY.  
FT DISULFID 46 74 BY SIMILARITY.  
FT DISULFID 78 98 BY SIMILARITY.  
FT DISULFID 99 104 BY SIMILARITY.  
FT LIPID 119 119 GPI-ANCHOR (POTENTIAL).  
FT VARIANT 63 63 D -> G (IN LY-6E.1).  
FT VARIANT 106 106 V -> A (IN LY-6E.1).  
SQ SEQUENCE 134 AA; 14377 MW; 883C771EB3430315 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 134;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 916 LLQTL 921  
Db 129 LLQTL 134

RESULT 198  
RR9\_GUITH STANDARD; PRT; 134 AA.  
AC PI9459; O46916;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S9.  
GN RPS9.  
OS Guillardia theta (Cryptomonas phi).  
OC Chloroplast.  
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
OX NCBI\_TaxID=55529;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97283757; PubMed=9137835;  
RA Wang S.L., Liu X.-Q., Douglas S.E.;  
RT "The large ribosomal protein gene cluster of a cryptomonad plastid:  
RT gene organization, sequence and evolutionary implications.";  
RL Biochem. Mol. Biol. Int. 41:1035-1044(1997).  
RN [2]  
RP SEQUENCE OF 97-134 FROM N.A.  
RX MEDLINE=91330343; PubMed=1868578;  
RA Douglas S.E.;  
RT "Unusual organization of a ribosomal protein operon in the plastid  
RT genome of Cryptomonas phi: evolutionary considerations.";  
RL Curr. Genet. 19:289-294(1991).  
CC -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF041468; AAC35726.1; -  
DR InterPro: IPR000754; -  
DR Pfam: PF00380; Ribosomal\_S9; 1.  
DR PROSITE; PS00360; RIBOSOMAL\_S9; 1.  
KW Ribosomal protein; Chloroplast.  
SQ SEQUENCE 134 AA; 15078 MW; 4AE5AC7E203DA10D CRC64;

Query Match 0.5%; Score 6; DB 1; Length 134;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1061 GQVEAI 1066  
Db 78 GQVEAI 83

RESULT 199  
VA2\_BPT5 STANDARD; PRT; 134 AA.  
AC F23541;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 01-NOV-1991 (Rel. 20, Last annotation update)  
DE A2 PROTEIN.  
OS Bacteriophage T5.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.  
OX NCBI\_TaxID=10726;  
RN [1]

```
RP SEQUENCE.
RX MEDLINE=91282771; Pubmed=2059212;
RA Snyder C.E. Jr.;
RT "Amino acid sequence of the bacteriophage T5 gene A2 protein.";
RL Biochem. Biophys. Res. Commun. 177:1240-1246(1991).
RN [2]
RP SEQUENCE OF 1-28.
RX MEDLINE=82256600; Pubmed=7049170;
RA Fox J.W., Barish A., Snyder C.E. Jr., Benzinger R.;
RT "Amino terminal sequence of the bacteriophage T5-coded gene A2
protein.";
RL Biochem. Biophys. Res. Commun. 106:265-269(1982).
CC -!- FUNCTION: MAY COMPETE WITH THE HOST OMFP PROTEIN FOR BINDING TO
LIPOLYSACCHARIDES.
CC -!- MISCELLANEOUS: IT IS ONE OF THE TWO PROTEINS ENCODED BY FIRST-STEP
TRANSFER DNA.
CC -!- SIMILARITY: STRONG, TO THE PRODUCT OF GENE A2-A3 OF BACTERIOPHAGE
BF23.
CC PIR: JQ1034; WXBPT5.
DR DNA-binding.
KW
SQ SEQUENCE 134 AA; 14200 MW; AD2BB0EC050F45F3 CRC64;

Query Match 0.5%; Score 6; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1008 ASLESA 1013
DB 105 ASLESA 110
|||||

RESULT 200
Y274_BUCAI
ID Y274_BUCAI STANDARD; PRT; 135 AA.
AC P57362;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PUTATIVE ACYL-COA THIOESTER HYDROLASE BU274 (EC 3.1.2.-).
GN BU274.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOKYO 1998;
RX MEDLINE=20445173; Pubmed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- SIMILARITY: BELONGS TO THE ACYL COENZYME A HYDROLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001118; BABI2984.1;
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 135 AA; 14724 MW; 6815A013965CBC8F CRC64;

Query Match 0.5%; Score 6; DB 1; Length 135;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 689 GKKVAT 694
```

Db 51 GKKVAT 56

|||||

Search completed: August 29, 2001, 09:38:38  
Job time: 249 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2001, 09:33:09 ; Search time 27.64 Seconds  
(without alignments)  
2842.572 Million cell updates/sec

Title: US-09-360-934A-3

Perfect score: 1296

Sequence: 1 MEIQOQTHRKINRPLVSLALV.....HNLISNIGHFASNLGMRYSF 1296

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 200 summaries.

Database :

A\_Geneseq\_0601.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	794	61.3	1296	14	AA411198
2	110	8.5	1288	18	AAW55347
3	110	8.5	1288	18	AAW55685
4	99	7.6	1287	16	AAW79944
5	95	7.3	176	18	AAW55346
6	87	6.7	326	21	AAW52500
7	87	6.7	578	21	AAW52575
8	87	6.7	1290	19	AAW98269
9	82	6.3	92	21	AAW52604
10	71	5.5	407	18	AAW55290
11	71	5.5	407	18	AAW20652

12	59	4.6	90	18	AAW20321	H. pylori protein.
13	59	4.6	216	20	AAW89963	Antigen from clust
14	48	3.7	288	18	AAW27720	H. pylori Vaca pro
15	45	3.5	224	21	AAW52506	Helicobacter pylor
16	37	2.9	141	18	AAW20260	H. pylori protein.
17	37	2.9	141	18	AAW24621	H. pylori protein.
18	37	2.9	211	18	AAW53324	H. pylori ORF 07ap
19	37	2.9	211	18	AAW20755	H. pylori ORF 07ap
20	37	2.9	257	18	AAW55462	H. pylori ORF 07ap
21	34	2.6	133	18	AAW55437	H. pylori ORF hp5p
22	29	2.2	63	18	AAW20382	H. pylori derived
23	23	1.8	23	14	AAW41353	N-terminal sequenc
24	20	1.5	93	18	AAW20362	H. pylori protein.
25	20	1.5	93	18	AAW55322	H. pylori ORF 07ap
26	20	1.5	95	18	AAW20754	H. pylori protein.
27	20	1.5	513	16	AAW79945	Helicobacter pylor
28	13	1.0	16	14	AAW41350	Peptide fragment o
29	10	0.8	15	18	AAW20215	H. pylori derived
30	8	0.6	246	18	AAW20825	H. pylori cytoplas
31	31	0.6	246	19	AAW11003	H. pylori ORF hp2e
32	8	0.6	293	19	AAW80647	S. pneumoniae tran
33	8	0.6	325	18	AAW20372	H. pylori cytoplas
34	8	0.6	325	18	AAW24644	H. pylori cytoplas
35	8	0.6	1213	18	AAW55735	H. pylori ORF 07ee
36	8	0.6	1851	22	AAU00023	Human activated T-
37	8	0.6	2893	19	AAW98828	H. pylori GHPO 148
38	8	0.6	2893	19	AAW71556	Helicobacter polyp
39	8	0.6	2902	22	AAW46351	H. pylori HPN165 p
40	7	0.5	8	14	AAW41351	Peptide fragment o
41	7	0.5	13	16	AAW84029	Murine MHC class I
42	7	0.5	13	16	AAW84032	Murine MHC class I
43	7	0.5	16	16	AAW84035	Murine MHC class I
44	7	0.5	16	20	AAW87824	Epitope of a Bcl-2
45	7	0.5	16	22	AAW74141	Bax epitope #11.
46	7	0.5	17	16	AAW82531	MHC groove specifi
47	7	0.5	17	18	AAW34191	Bt-E.alpha. Homo
48	7	0.5	17	21	AAW79534	Bt-E.alpha+Y. Hom
49	7	0.5	17	22	AAW48952	Mouse I-E molecule
50	7	0.5	17	22	AAW41352	I-Ab-restricted MH
51	7	0.5	20	14	AAW71467	Peptide fragment o
52	7	0.5	22	21	AAW90740	Human BCL-2 amino
53	7	0.5	25	17	AAW87005	Class II MHC anti
54	7	0.5	25	18	AAW21833	Rat RT1.Dalpha cla
55	7	0.5	57	21	AAW20215	Human secreted pro
56	7	0.5	62	16	AAW71467	Mature thermophil
57	7	0.5	67	16	AAW82540	Hybrid IA beta cha
58	7	0.5	68	22	AAW89555	DICE-II I-Ab-restr
59	7	0.5	72	21	AAW16568	Arabidopsis thalia
60	7	0.5	72	21	AAW52897	Arabidopsis thalia
61	7	0.5	81	18	AAW10492	Alphal region of C
62	7	0.5	85	16	AAW71469	Premature thermoph
63	7	0.5	101	21	AAW40396	Human ORFX ORF160
64	7	0.5	102	20	AAW89830	N-terminal sequenc
65	7	0.5	118	21	AAW00269	Human secreted pro
66	7	0.5	129	21	AAW03754	Human secreted pro
67	7	0.5	133	21	AAW54190	Human pancreatic c
68	7	0.5	134	18	AAW55658	H. pylori ORF 11ae
69	7	0.5	145	21	AAW00183	Human secreted pro
70	7	0.5	152	18	AAW55229	Human secreted pro
71	7	0.5	160	20	AAW59700	Secreted protein 7
72	7	0.5	160	21	AAW23612	Human secreted pro
73	7	0.5	160	21	AAW87357	Human signal pepti
74	7	0.5	161	20	AAW27684	Pinus radiata cell
75	7	0.5	173	21	AAW25422	Peptide Seq ID No:
76	7	0.5	174	20	AAW06802	Peptide Seq ID No:
77	7	0.5	174	20	AAW06803	Human secreted pro
78	7	0.5	176	21	AAW00963	Soluble fused MHC
79	7	0.5	207	18	AAW10493	H. pylori inner me
80	7	0.5	207	18	AAW20276	H. pylori cell env
81	7	0.5	208	18	AAW20824	Arabidopsis thalia
82	7	0.5	210	21	AAW40598	Human colon cancer
83	7	0.5	214	21	AAW53450	Gene 28 human secr
84	7	0.5	224	21	AAW51810	



85	7	0.5	228	22	AAB69576	Swine leukocyte an
86	7	0.5	228	22	AAB69577	Swine leukocyte an
87	7	0.5	228	22	AAB69578	Swine leukocyte an
88	7	0.5	228	22	AAB69579	Swine leukocyte an
89	7	0.5	228	22	AAB69580	Swine leukocyte an
90	7	0.5	228	22	AAB69581	Swine leukocyte an
91	7	0.5	228	22	AAB69582	Swine leukocyte an
92	7	0.5	229	16	AAW74037	MHC polypeptide HL
93	7	0.5	229	17	AAW01021	Apoptosis-BLOCKING
94	7	0.5	229	17	AAW94348	Human Bcl-2 mutant
95	7	0.5	232	17	AAW01019	Apoptosis-BLOCKING
96	7	0.5	232	17	AAW01020	Apoptosis-BLOCKING
97	7	0.5	232	20	AAW94346	Human Bcl-2 mutant
98	7	0.5	232	20	AAW94347	Human Bcl-2 mutant
99	7	0.5	236	22	AAB35131	Murine Bcl-2. Mus
100	7	0.5	239	9	AAW80967	Sequence of bcl-2-
101	7	0.5	239	14	AAW42312	Bcl-2 oncogene pro
102	7	0.5	239	15	AAW47344	Human oncogene bcl
103	7	0.5	239	16	AAW70331	Human bcl-2 alpha
104	7	0.5	239	16	AAW71404	Human bcl-2 alpha
105	7	0.5	239	17	AAW02383	Human BCL2. Homo
106	7	0.5	239	17	AAW01018	Apoptosis-BLOCKING
107	7	0.5	239	19	AAW40217	Human bcl-2. Homo
108	7	0.5	239	20	AAW94345	Human Bcl-2 wild-t
109	7	0.5	239	20	AAW87810	A human Bcl-2 prot
110	7	0.5	239	20	AAW87812	A human Bcl-2 alph
111	7	0.5	239	21	AAW69203	Amino acid sequenc
112	7	0.5	239	22	AAW74127	Human bcl-2. Homo
113	7	0.5	239	22	AAW874129	Human bcl-2 alpha
114	7	0.5	239	22	AAW48288	Human Bcl-2 protei
115	7	0.5	239	22	AAW35130	Human Bcl-2. Homo
116	7	0.5	239	22	AAW50537	Human Bcl-2. Homo
117	7	0.5	245	19	AAW46943	Human Bcl-2 protei
118	7	0.5	245	20	AAW31652	HLA-DR2 alpha-Fos
119	7	0.5	248	18	AAW37341	DR alpha-DAF chime
120	7	0.5	251	19	AAW98711	H. pylori ORF 771
121	7	0.5	253	21	AAW68277	Class II alpha cha
122	7	0.5	253	21	AAW52931	Class II alpha cha
123	7	0.5	253	22	AAW58692	Class II alpha cha
124	7	0.5	256	21	AAW68282	Class II alpha cha
125	7	0.5	256	21	AAW52936	Class II alpha cha
126	7	0.5	256	22	AAW58697	Class II alpha cha
127	7	0.5	272	19	AAW21120	Human bcl2 proto-o
128	7	0.5	281	21	AAW58160	Lung cancer associ
129	7	0.5	283	8	AAW70392	Alpha-subunit of S
130	7	0.5	283	8	AAW70427	Sequence encoding
131	7	0.5	290	21	AAW32503	S. lavendulae mit
132	7	0.5	293	19	AAW86120	S. pneumoniae deri
133	7	0.5	294	19	AAW10990	H. pylori ORF 06p
134	7	0.5	297	18	AAW20738	H. pylori cytoplasm
135	7	0.5	298	16	AAW82538	Hydrid IA beta cha
136	7	0.5	299	18	AAW55450	H. pylori ORF 02ae
137	7	0.5	299	19	AAW98322	H. pylori GHPO 136
138	7	0.5	299	19	AAW73034	Helicobacter pylori
139	7	0.5	299	20	AAW89829	Protein encoded by
140	7	0.5	299	20	AAW89849	Antigen from clust
141	7	0.5	305	20	AAW00875	Synapocjanin isofo
142	7	0.5	316	21	AAW06505	Arabidopsis thalia
143	7	0.5	317	21	AAW40408	Arabidopsis thalia
144	7	0.5	323	21	AAW06504	Arabidopsis thalia
145	7	0.5	330	22	AAW67352	Protein encoded by
146	7	0.5	340	18	AAW55731	H. pylori ORF 14cp
147	7	0.5	340	21	AAW58310	Lung cancer associ
148	7	0.5	344	21	AAW15928	E. coli proliferat
149	7	0.5	351	21	AAW75482	Neisseria meningit
150	7	0.5	351	21	AAW75483	Neisseria meningit
151	7	0.5	354	21	AAW40467	Arabidopsis thalia
152	7	0.5	363	21	AAW06503	Arabidopsis thalia
153	7	0.5	365	21	AAW40466	Arabidopsis thalia
154	7	0.5	377	8	AAW70457	Sequence of gpc en
155	7	0.5	378	19	AAW60766	Murine Lunatic fri
156	7	0.5	380	21	AAW32503	Arabidopsis thalia
157	7	0.5	381	21	AAW21935	Arabidopsis thalia

158	7	0.5	387	19	AAW98493	H. pylori GHPO 107
159	7	0.5	387	20	AAW17202	H. pylori Outer me
160	7	0.5	387	21	AAW21934	Arabidopsis thalia
161	7	0.5	395	19	AAW98760	H. pylori GHPO 108
162	7	0.5	421	22	AAW76632	Cornebacterium q1
163	7	0.5	438	21	AAW40913	Zee mays protein f
164	7	0.5	443	21	AAW21037	Human nucleic acid
165	7	0.5	443	21	AAW87780	Human ADA2 protein
166	7	0.5	446	21	AAW09862	Arabidopsis thalia
167	7	0.5	446	21	AAW32502	Arabidopsis thalia
168	7	0.5	479	20	AAW31654	HLA-DR2 alpha-Fos-
169	7	0.5	495	21	AAW25739	Arabidopsis thalia
170	7	0.5	495	21	AAW37847	Arabidopsis thalia
171	7	0.5	495	21	AAW39756	Arabidopsis thalia
172	7	0.5	498	21	AAW40912	Zee mays protein f
173	7	0.5	510	12	AAW15354	Protein deduced fr
174	7	0.5	510	19	AAW79740	Soybean wild-type
175	7	0.5	510	19	AAW79741	Soybean mutant myo
176	7	0.5	510	20	AAW96259	Phyate protein.
177	7	0.5	510	20	AAW97882	Maize myo-inositol
178	7	0.5	510	21	AAW50575	Arabidopsis thalia
179	7	0.5	510	22	AAW48935	Brassica napus myo
180	7	0.5	511	21	AAW09861	Arabidopsis thalia
181	7	0.5	511	21	AAW32501	Arabidopsis thalia
182	7	0.5	534	21	AAW09860	Arabidopsis thalia
183	7	0.5	536	20	AAW24477	Arabidopsis thalia
184	7	0.5	539	18	AAW41608	Nicotiana panicula
185	7	0.5	539	18	AAW41609	Soybean protox-1.
186	7	0.5	539	18	AAW25740	Cotton protox-1.
187	7	0.5	539	22	AAW72907	Cotton protox-1.
188	7	0.5	539	22	AAW72921	Cotton protox-1.
189	7	0.5	539	22	AAW72922	Cotton protox-1.
190	7	0.5	564	19	AAW74581	5' fragment of mem
191	7	0.5	581	21	AAW50574	Arabidopsis thalia
192	7	0.5	592	21	AAW25738	Arabidopsis thalia
193	7	0.5	592	21	AAW37846	Arabidopsis thalia
194	7	0.5	592	21	AAW39755	Arabidopsis thalia
195	7	0.5	594	19	AAW74579	Arabidopsis thalia
196	7	0.5	594	19	AAW74580	Rabbit membrane pr
197	7	0.5	612	20	AAW31655	Human membrane pr
198	7	0.5	628	21	AAW81659	HLA-DR2 alpha-Fos-
199	7	0.5	645	21	AAW50573	Streptococcus pneu
200	7	0.5	673	21	AAW25737	Arabidopsis thalia

## ALIGNMENTS

RESULT	1
ID	AAW41198 standard; Protein; 1296 AA.
AC	AAW41198;
XX	
DT	17-MAR-1994 (first entry)
DE	
CT	
KW	Cytotoxin; CT; H. pylori; precursor; vacuolation; cell death; hsp60;
KW	heat shock protein; cytotoxin-associated immunodominant antigen; CAT;
KW	hsp; type B; gastritis; peptic ulcer; eukaryote; gastric tumours.
XX	
OS	Helicobacter pylori.
XX	
FT	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	1..33
FT	/note="Signal peptide"
FT	34..1296
FT	/note="Mature CT"
PN	WO9318150-A.
PD	16-SEP-1993.

XX 02-MAR-1993; 93WO-EP00472.  
 XX 02-MAR-1992; 92IT-OFI0052.  
 PR 25-JAN-1993; 93WO-EP00158.  
 XX (BIOC-) BIOCINE SCLAVO SPA.  
 XX Bugnoli M, Covacci A, Macchia G, Rappuoli R, Telford J;  
 XX WPI; 1993-303464/38.  
 DR N-PSDB; AAQ48732.  
 XX Recombinant Helicobacter pylori protein and corresp. gene - is a  
 PT cytotoxin, antigen or heat shock protein used for treating and  
 PT preventing type B gastritis, gastric ulcers and gastric tumours  
 XX Claim 3; Fig 2; 83pp; English.  
 XX This sequence represents the precursor form of cytotoxin, CT, of H.  
 CC pylori. The DNA encoding this protein was isolated from two  
 CC overlapping fragments corresponding to about 10 kb of the H. pylori  
 CC genome. This protein has a molecular weight of 139.8 kD and serves as  
 CC a precursor to a protein having an approximate weight of 100 kD and a  
 CC cytotoxic activity. The amino acid positions 34-56 indicate a  
 CC previously isolated N-terminal peptide. This sequence is preceded by  
 CC 33 amino acids which resemble prokaryotic leader sequences, thus this  
 CC sequence is likely to represent the N-terminal of the mature protein.  
 CC A previously isolated cytotoxin of 87 kD may be produced by further  
 CC processing or proteolytic degradation during purification. The mature  
 CC cytotoxin causes vacuolation and death of a number of eukaryotic cell  
 CC types. This protein, and others derived from H. pylori, esp.  
 CC cytotoxin-associated immunodominant (CAI) antigen or a heat shock  
 CC protein (hsp) (see also AA41199-200), may be used to treat, prevent and  
 CC diagnose H. pylori infection. H. pylori is the causative agent of type  
 CC B gastritis, peptic ulcers and gastric tumours.  
 XX Sequence 1296 AA;

Query Match 61.3%; Score 794; DB 14; Length 1296;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 894; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 206 RVGSAGKASTVLTQASGITSKNAEISLYDGATNLASSSVKLMGNVWMLQYV 265  
 Db |||||  
 QY 206 RVGSAGKASTVLTQASGITSKNAEISLYDGATNLASSSVKLMGNVWMLQYV 265  
 Db |||||  
 QY 266 GAYLAPSYSTINTSKVTGEVNFHILTVGDKNAQAQIITANKTNTGTLDLWQSAGLNIIA 325  
 Db |||||  
 QY 266 GAYLAPSYSTINTSKVTGEVNFHILTVGDKNAQAQIITANKTNTGTLDLWQSAGLNIIA 325  
 Db |||||  
 QY 326 PPEGYKDKPNTPSQSAKNDKNEAKNDKQESSQNNSTQVINPPNSAOKTEVQPTQV 385  
 Db |||||  
 QY 326 PPEGYKDKPNTPSQSAKNDKNEAKNDKQESSQNNSTQVINPPNSAOKTEVQPTQV 385  
 Db |||||  
 QY 386 IDGPPAGKQDVNINRINTNADGTRVGGFKASLTNAHLHIGKGVNLSNQASGRSL 445  
 Db |||||  
 QY 386 IDGPPAGKQDVNINRINTNADGTRVGGFKASLTNAHLHIGKGVNLSNQASGRSL 445  
 Db |||||  
 QY 446 IVENLTGNTVDGPLRVNNOVGGYALAGSSANFEKAGTDKNGTATFNNDISLGRFVNL 505  
 Db |||||  
 QY 446 IVENLTGNTVDGPLRVNNOVGGYALAGSSANFEKAGTDKNGTATFNNDISLGRFVNL 505  
 Db |||||  
 QY 506 KVDAHTANFKGIDTCNGGFNTLDFSGVTGDKVNIKNLITASTNVAVKNFNINELIVKTNIGI 565  
 Db |||||  
 QY 506 KVDAHTANFKGIDTCNGGFNTLDFSGVTGDKVNIKNLITASTNVAVKNFNINELIVKTNIGI 565  
 Db |||||  
 QY 566 SVGEYTHFSEDIGSOSRINTVRLTGTSLFSGGVKFKGKGLVIDEYFSPWNYFDARN 625  
 Db |||||  
 QY 566 SVGEYTHFSEDIGSOSRINTVRLTGTSLFSGGVKFKGKGLVIDEYFSPWNYFDARN 625  
 Db |||||  
 QY 626 IKNVEITNKLAFGPGSGPWGTSKLMFNLTLCQNAVMDYSOFNSLTIOGDPFINNOGTINY 685  
 Db |||||

Db 626 IKNVEITNKLAFGPGSGPWGTSKLMFNLTLCQNAVMDYSQFINLTIGDFINQGTINY 685  
 QY 686 LVRGGKAVATLSVGNAAAMFNNDIDSATGFYKPLIKINSAQDLIKNTHVLLKAKIIGY 745  
 Db 686 LVRGGKAVATLSVGNAAAMFNNDIDSATGFYKPLIKINSAQDLIKNTHVLLKAKIIGY 745  
 QY 746 NVSTGTNGISNVNLEQFKERLALYNNNNRMDTCVYVRNTDDIKACGMAIGDQSMVNNPDN 805  
 Db 746 NVSTGTNGISNVNLEQFKERLALYNNNNRMDTCVYVRNTDDIKACGMAIGDQSMVNNPDN 805  
 QY 806 YKYLICKAKWNGIKSTANGSKISVYILGNSTPTENGNTNLTPTNTTNSARANNALAQ 865  
 Db 806 YKYLICKAKWNGIKSTANGSKISVYILGNSTPTENGNTNLTPTNTTNSARANNALAQ 865  
 QY 866 NAPFAQPSATPNLVAIQHDFGTIESVFELANRSKIDITLYANSAGQGRDLOTLIDSH 925  
 Db 866 NAPFAQPSATPNLVAIQHDFGTIESVFELANRSKIDITLYANSAGQGRDLOTLIDSH 925  
 QY 926 DAGYARKMIDAYSANEITKOLNTATTTLNNIASLEHKTSGLOTLSLSNAMILNSRLVNL 985  
 Db 926 DAGYARKMIDAYSANEITKOLNTATTTLNNIASLEHKTSGLOTLSLSNAMILNSRLVNL 985  
 QY 986 RRTHNHIDSFARLQALQKQKFALESAAEVLYQFAPKYEKPTNVWANAIGGTSLNNGSN 1045  
 Db 986 RRTHNHIDSFARLQALQKQKFALESAAEVLYQFAPKYEKPTNVWANAIGGTSLNNGSN 1045  
 QY 1046 ASLYGTSGAGVAYLNGOEAIVGGSGSYCYSSFNPNRANSLNSCANNTFGVYSRI 1100  
 Db 1046 ASLYGTSGAGVAYLNGOEAIVGGSGSYCYSSFNPNRANSLNSCANNTFGVYSRI 1100

RESULT 2  
 AAW55547  
 ID AAW55547 standard; Protein; 1288 AA.  
 XX AC AAW55547;  
 XX DT 24-JUN-1998 (first entry)  
 XX DE H. pylori ORF 14ee1924\_2458267\_c2\_93 secreted protein.  
 XX KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.  
 OS Helicobacter pylori.  
 XX PN W09737044-A1.  
 XX PD 09-OCT-1997.  
 XX PF 27-MAR-1997; 97WO-US05223.  
 XX PR 06-DEC-1996; 96US-0761318.  
 PR 29-MAR-1996; 96US-0625811.  
 PR 02-APR-1996; 96US-0758731.  
 PR 25-OCT-1996; 96US-0736905.  
 PR 28-OCT-1996; 96US-0738859.  
 XX (ASTR ) ASTRA AB.  
 XX PA Alm RA, Smith D;  
 XX PI WPI; 1997-503122/46.  
 XX DR N-PSDB; AAV24956.  
 XX DR Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX PS Disclosure; Page 750-753; 1145pp; English.  
 XX

CC This sequence is a *H. pylori* secreted protein. The protein may be used  
 CC in a vaccine to prevent or treat *H. pylori* infection or to identify  
 CC *H. pylori*-polypeptide binding compounds, useful as potential *H. pylori*  
 CC life cycle activators or inhibitors. The DNA and probes derived from it  
 CC may be used for the identification of *H. pylori* in a sample and the  
 CC diagnosis of *H. pylori* infection. Nucleic acid sequences complementary  
 CC to the DNA act as antisense sequences and can be used to prevent the  
 CC translation of *H. pylori* mRNA. Antibodies against the protein can be  
 CC used in immunoassays to evaluate the abundance and distribution of  
 CC *H. pylori*-specific antigens. The genomic sequence of *H. pylori*  
 CC (ATCC 55679) was determined from overlapping contigs generated by  
 CC mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely *H. pylori* antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from *H. pylori* by PCR  
 CC amplification for recombinant polypeptide production, e.g. in *E. coli*  
 CC hosts.

XX Sequence 1288 AA;

Query Match 8.5%; Score 110; DB 18; Length 1288;  
 Best Local Similarity 100.0%; Pred. No. 4e-101;  
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 DTKNGTATFNNDISLGRFVNLKVDHTANFKGIDTNGGNTLDFSGVTDKVNINKLITA 544  
 Db |||||||  
 QY 478 dtkngtatfnndisigrfvlkvdahtanfkgidtgngfntldfsgvtdkvninklita 537  
 Db |||||||  
 QY 545 STNVAVKNENINELIVKTNIGSVGEYTHFSEDIGSQSRINTVRLTGTRS 594  
 Db |||||||  
 QY 538 stnvavknfninellivktnigsvgeythfse digsqsrintvrltgtrs 587

RESULT 3  
 AAW55685

ID AAW55685 standard; Protein: 1288 AA.

XX AAW55685;

XX 07-JUL-1998 (first entry)

DE H. Pylori ORF 07ee11402\_2458267\_c3\_108 cell envelope OMP.

XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.

OS Helicobacter pylori.

XX WO9737044-A1.

XX 09-OCT-1997.

XX 27-MAR-1997; 97WO-US05223.

XX 06-DEC-1996; 96US-0761318.

XX 29-MAR-1996; 96US-0625811.

XX 02-APR-1996; 96US-0738731.

XX 25-OCT-1996; 96US-0736903.

XX 28-OCT-1996; 96US-0738859.

XX (ASTR ) ASTRA AB.

XX Alm RA, Smith D;

XX WPI; 1997-503122/46.

XX N-PSDB; AAV25094.

XX Helicobacter pylori nucleic acid sequences and encoded

PT polypeptide(s) - useful in vaccines to treat or prevent *H. pylori*  
 XX infection and for diagnosis of *H. pylori* infection  
 XX Disclosure; Pages 939-942; 1145pp; English.

CC This sequence is a *H. pylori* cell envelope outer membrane (OMP)  
 CC protein having a terminal Phe residue.

CC The protein may be used in a vaccine to prevent or treat *H. pylori*  
 CC infection or to identify *H. pylori* polypeptide binding compounds, The  
 CC useful as potential *H. pylori* life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC *H. pylori* in a sample and the diagnosis of *H. pylori* infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of *H. pylori* mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of *H. pylori*-specific antigens. The genomic sequence of  
 CC *H. pylori* (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely *H. pylori* antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from *H. pylori* by PCR  
 CC amplification for recombinant polypeptide production, e.g. in *E. coli*  
 CC hosts.

XX Sequence 1288 AA;

Query Match 8.5%; Score 110; DB 18; Length 1288;  
 Best Local Similarity 100.0%; Pred. No. 4e-101;  
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 DTKNGTATFNNDISLGRFVNLKVDHTANFKGIDTNGGNTLDFSGVTDKVNINKLITA 544  
 Db |||||||  
 QY 478 dtkngtatfnndisigrfvlkvdahtanfkgidtgngfntldfsgvtdkvninklita 537  
 Db |||||||  
 QY 545 STNVAVKNENINELIVKTNIGSVGEYTHFSEDIGSQSRINTVRLTGTRS 594  
 Db |||||||  
 QY 538 stnvavknfninellivktnigsvgeythfse digsqsrintvrltgtrs 587

RESULT 4

AAAR79944

ID AAR79944 standard; Protein: 1287 AA.

XX AAR79944;

XX 26-MAR-1996 (first entry)

DE Helicobacter pylori vacuolating toxin.

XX Vacuolating toxin; vaccine; immunisation; therapy; mutant;  
 KW infection; Helicobacter pylori.

OS Helicobacter pylori.

XX WO9522988-A1.

XX 31-AUG-1995.

XX 23-FEB-1995; 95WO-US02219.

XX 23-FEB-1994; 94US-0200232.

XX (UYVA-) UNIV VANDERBILT.

XX Blaser MJ; Cover TL;

XX WPI; 1995-311383/40.

XX N-PSDB; AAT04132.

PT Isolated DNA encoding Helicobacter pylori vacuolating toxin - useful  
 PT for immunisation against H. pylori infection  
 XX  
 PS Claim 7; Page 43-47; 64pp; English.

CC The nucleic acid encoding the Helicobacter pylori vacuolating toxin  
 CC and a genetically altered mutant strain of H. pylori which contains  
 CC a foreign nucleic acid and does not express a functional vacuolating  
 CC toxin may be used to immunise a subject against H. pylori infection.  
 CC They may possibly also be used therapeutically.

XX Sequence 1287 AA;

SQ Query Match 7.6%; Score 99; DB 16; Length 1287;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-90;  
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 VGNAAAMFNNDIDSATGFKPLIKINSADLIKNTHEVLLKAKIIGYGNVSTGTNGISN 756  
 Db 689 vgnaaamfnndidsatgfkplikinsadlikntehvllkakiigynvstgtngisn 748  
 QY 757 VNLEEFKRLALYNNNRMDTCVVRNTDDIKACGMAIG 795  
 Db 749 vnleedfkerlalynnnrmdtcvvrntddikacgmaig 787

RESULT 5  
 AAW55346  
 ID AAW55346 standard; Protein; 176 AA.  
 AC AAW55346;  
 XX  
 XX 17-JUN-1998 (first entry)

DE H. pylori ORF 07ep30818orf4 protein.

KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.  
 XX Helicobacter pylori.

OS WO9737044-A1.

PN '09-OCT-1997.

PD 27-MAR-1997; 97WO-US05223.

PF 06-DEC-1996; 96US-0761318.

PR 29-MAR-1996; 96US-0625811.

PR 02-APR-1996; 96US-0758731.

PR 25-OCT-1996; 96US-0736905.

PR 28-OCT-1996; 96US-0738859.

XX (ASTR ) ASTRA AB.

XX Alm RA, Smith D;

XX WPI; 1997-503122/46.

XX N-PSDB; AAV24755.

PT Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX  
 PS Claim 14; Page 570-571; 1145pp; English.

CC This sequence is a H. pylori protein of unspecified function.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of

CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunosays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.

SQ Sequence 176 AA;

Query Match 7.3%; Score 95; DB 18; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-87;  
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 701 AAMFNNDIDSATGFKPLIKINSADLIKNTHEVLLKAKIIGYGNVSTGTNGISNVNLE 760  
 Db 1 aamfnndidsatgfkplikinsadlikntehvllkakiigynvstgtngisnvnle 60

QY 761 EOPKRLALYNNNRMDTCVVRNTDDIKACGMAIG 795

Db 61 eqfkerlalynnnrmdtcvvrntddikacgmaig 95

RESULT 6

AAB52500

ID AAB52500 standard; Protein; 326 AA.

AC AAB52500;

DT 23-FEB-2001 (first entry)

DE Helicobacter pylori bait polypeptide #18.

KW Helicobacter pylori; two-hybrid system; protein-protein interaction;  
 KW bait polypeptide; gastric ulcer; antibacterial.

OS Helicobacter pylori.

PN WO200066722-A1.

PD 09-NOV-2000.

PF 14-APR-2000; 2000WO-IB00603.

PR 30-APR-1999; 99EP-0401066.

XX (HYBR-) HYBRIGENICS SA.

XX Legrain P, Selig L, Rain J;

XX WPI; 2000-687535/67.

XX N-PSDB; AAC97246.

XX A two-hybrid system for identifying compounds useful in the treatment  
 of e.g. gastric ulcers comprises producing a collection of recombinant  
 cell clones -

PS Example 5; Page 101; 267pp; English.

CC The present sequence is a bait polypeptide used in a Helicobacter  
 CC pylori two-hybrid screen to identify protein-protein interactions.  
 CC The method is used to identify a recombinant cell clone expressing a  
 CC prey polypeptide which is capable of interacting with the bait  
 CC polypeptide. The two hybrid system is useful for screening compounds

CC for antibacterial activity. It may be used in the treatment of gastric  
 CC ulcers. The polynucleotides are useful as amplification primers or  
 CC specific detection probes. The polypeptides, vectors or host cells can  
 CC be used as immunogens to produce mono- or polyclonal antibodies. The  
 CC polynucleotides, polypeptides, antibodies, vectors, host cells or  
 CC modulating agents can be used to produce a pharmaceutical composition.  
 XX  
 SQ Sequence 326 AA;

Query Match 6.7%; Score 87; DB 21; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-78;  
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTGNITVDGRLRVNNQVGGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 506  
 |||||  
 DB 236 venltgntitvdgplrvnnqvggvalagssanfeakgttdkngtatfnndisigrfvnlk 295  
 |||||  
 QY 507 VDAHTANFKGIDTNGGNTLDFSGVT 533  
 |||||  
 DB 296 vdahtanfkgtidnggntldfsgvt 322  
 |||||

## RESULT 7

AAB52575 standard; Protein; 578 AA.

XX  
 AC AAB52575;  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Helicobacter pylori bait polypeptide #93.

KW Helicobacter pylori; two-hybrid system; protein-protein interaction;  
 KW bait polypeptide; gastric ulcer; antibacterial.

XX Helicobacter pylori.

XX WO200066722-A1.

XX 09-NOV-2000.

PF 14-APR-2000; 2000WO-IB00603.

PR 30-APR-1999; 99EP-0401066.

PA (HYBR-) HYBRIGENICS SA.

PI Legrain P, Sellig L, Rain J;

DR WPI; 2000-687535/67.

DR N-PSDB; AAC97321.

XX  
 PT A two-hybrid system for identifying compounds useful in the treatment  
 PT of e.g. gastric ulcers comprises producing a collection of recombinant  
 PT cell clones.

XX Example 5; Page 202-204; 267pp; English.

XX The present sequence is a bait polypeptide used in a Helicobacter  
 CC pylori two-hybrid screen to identify protein-protein interactions.  
 CC The method is used to identify a recombinant cell clone expressing a  
 CC prey polypeptide which is capable of interacting with the bait  
 CC polypeptide. The two hybrid system is useful for screening compounds  
 CC for antibacterial activity. It may be used in the treatment of gastric  
 CC ulcers. The polynucleotides are useful as amplification primers or  
 CC specific detection probes. The polypeptides, vectors or host cells can  
 CC be used as immunogens to produce mono- or polyclonal antibodies. The  
 CC polynucleotides, polypeptides, antibodies, vectors, host cells or  
 CC modulating agents can be used to produce a pharmaceutical composition.

XX Sequence 578 AA;

Query Match 6.7%; Score 87; DB 21; Length 578;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-78;  
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTGNITVDGRLRVNNQVGGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 506  
 |||||  
 DB 261 venltgntitvdgplrvnnqvggvalagssanfeakgttdkngtatfnndisigrfvnlk 320  
 |||||

QY 507 VDAHTANFKGIDTNGGNTLDFSGVT 533  
 |||||

DB 321 vdahtanfkgtidnggntldfsgvt 347  
 |||||

## RESULT 8

AAW98269 standard; Protein; 1290 AA.

XX  
 AC AAW98269;

DT 31-MAR-1999 (first entry)

XX H. pylori GHPO 374 protein.

DE H. pylori GHPO 374 protein.

XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.

XX Helicobacter pylori.

XX WO9843478-A1.

PD 08-OCT-1998.

XX 01-APR-1998; 98WO-US06371.

PR 29-JUL-1997; 97US-0902615.

PR 01-APR-1997; 97US-0833457.

PR 24-JUN-1997; 97US-0881227.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;

DR WPI; 1998-542293/46.

DR N-PSDB; AAX13988.

XX New isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 PT infections and gastrointestinal diseases

PS Claim 8; Page 273-278; 2054pp; English.

XX This sequence represents a Helicobacter pylori GHPO protein of the  
 CC invention. The polypeptides can be used for preventing or treating  
 CC Helicobacter infections, and gastroduodenal diseases associated with  
 CC these infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
 CC used for the production of antibodies. The products can also be for  
 CC detection and diagnosis.

XX Sequence 1290 AA;

Query Match 6.7%; Score 87; DB 19; Length 1290;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-78;  
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 VENLTGNITVDGRLRVNNQVGGYALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 506  
 |||||  
 DB 442 venltgntitvdgplrvnnqvggvalagssanfeakgttdkngtatfnndisigrfvnlk 501  
 |||||

QY 507 VDAHTANFKGIDTNGGNTLDFSGVT 533  
 |||||

Db 502 vdahtanfgldtgnngfntldfsgvt 528  
 |||||

RESULT 9  
 AAB52604  
 ID AAB52604 standard; Protein; 92 AA.  
 AC AAB52604;  
 XX  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Helicobacter pylori bait polypeptide #122.  
 XX  
 KW Helicobacter pylori; two-hybrid system; protein-protein interaction;  
 KW bait polypeptide; gastric ulcer; antibacterial.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO200066722-A1.  
 XX  
 PD 09-NOV-2000.  
 XX  
 PF 14-APR-2000; 2000WO-IB00603.  
 XX  
 PR 30-APR-1999; 99EP-0401066.  
 XX  
 PA (HYBR-) HYBRIGENICS SA.  
 XX  
 PI Legrain P, Selig L, Rain J;  
 XX  
 DR WPI; 2000-687535/67.  
 DR N-PSDB; AAC97350.  
 XX  
 PT A two-hybrid system for identifying compounds useful in the treatment  
 PT of e.g. gastric ulcers comprises producing a collection of recombinant  
 PT cell clones -  
 XX  
 PS Example 5; Page 234; 267pp; English.  
 XX  
 CC The present sequence is a bait polypeptide used in a Helicobacter  
 CC pylori two-hybrid screen to identify protein-protein interactions.  
 CC The method is used to identify a recombinant cell clone expressing a  
 CC prey polypeptide which is capable of interacting with the bait  
 CC polypeptide. The two hybrid system is useful for screening compounds  
 CC for antibacterial activity. It may be used in the treatment of gastric  
 CC ulcers. The polynucleotides are useful as amplification primers or  
 CC specific detection probes. The polypeptides, vectors or host cells can  
 CC be used as immunogens to produce mono- or polyclonal antibodies. The  
 CC polynucleotides, polypeptides, antibodies, vectors, host cells or  
 CC modulating agents can be used to produce a pharmaceutical composition.  
 XX  
 SQ Sequence 92 AA;  
 Query Match 6.3%; Score 82; DB 21; Length 92;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-74;  
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 452 GNITVDGFLRVNNOVGVALAGSSANFFKAGTDKNGTATFNNDISLGRFVNLKVDHT 511  
 Db 1 gnitvdgflrvnnvggvalagssanffkagtdkngtattfnndisigrfvlkvdht 60  
 OY 512 ANFKGIDTNGNGFNLDFSGVT 533  
 Db 61 anfkidtgnggfnldfsgvt 82  
 RESULT 10  
 AAW5290  
 ID AAW5290 standard; Protein; 407 AA.  
 XX  
 AC AAW5290;

XX  
 DT 15-JUN-1998 (first entry)  
 XX  
 DE H. pylori ORF 02ep30607orf19 protein.  
 XX  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9737044-A1.  
 XX  
 PD 09-OCT-1997.  
 XX  
 PF 27-MAR-1997; 97WO-US05223.  
 XX  
 PR 06-DEC-1996; 96US-0761318.  
 PR 29-MAR-1996; 96US-0625811.  
 PR 02-APR-1996; 96US-0758731.  
 PR 25-OCT-1996; 96US-0736905.  
 PR 28-OCT-1996; 96US-0738859.  
 XX  
 PA (ASTR ) ASTRA AB.  
 XX  
 PI Alm RA, Smith D;  
 XX  
 DR WPI; 1997-503122/46.  
 DR N-PSDB; AAV24699.  
 XX  
 PT Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX  
 PS Claim 14; Page 524-525; 1145pp; English.  
 XX  
 CC This sequence is a H. pylori protein of unspecified function.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.  
 XX  
 SQ Sequence 407 AA;

Query Match 5.5%; Score 71; DB 18; Length 407;  
 Best Local Similarity 100.0%; Pred. No. 2e-62;  
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 31 SHAAFTTVIIPAIIVGGIATGATGTVSGLLSWGLKQAEANKTKDPKDKVRIQAGKF 90  
 Db 47 shaafttviipaiivggiatgavgtvsgllswgkqaeanktkdpkdkvriqagkf 106  
 OY 91 NEFPNKEYDLY 101  
 Db 107 nefpnkeydly 117

## RESULT 11

AAW20652  
ID AAW20652 standard; protein; 407 AA.

XX AC  
XX AC AAW20652;

DT XX  
DE XX 14-JUL-1997 (first entry)

DE XX H. pylori derived protein.

XX KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;  
XX OS Helicobacter pylori.

XX PN WO9640893-A1.

XX PD 19-DEC-1996.

XX PF 06-JUN-1996; 96WO-US09122.

XX PR 01-APR-1996; 96US-0630405.

XX PR 07-JUN-1995; 95US-0487032.

XX PA (ASTR ) ASTRA AB.

XX PI Berglindh OT, Smith D, Mellgaard BL;

XX DR WPI; 1997-052306/05.

XX DR N-PSDB; AAT67905.

XX XX Helicobacter pylori nucleic acid sequences and related

PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PT infection, and to detect Helicobacter

XX PS Disclosure; Page 1074-1075; 1481pp; English.

XX CC The present sequence is a H. pylori derived protein of unspecified

CC function, no further details are given in the specification.

CC The protein may be used in a vaccine to prevent or treat H. pylori

CC infection or to identify H. pylori polypeptide binding compounds,

CC useful as potential H. pylori life cycle activators or inhibitors.

CC The genomic sequence of H. pylori (ATCC 55679) was determined from

CC overlapping contigs generated by mechanically shearing the bacterial

CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,

CC and the predicted coding regions defined by computer evaluation. To

CC identify likely H. pylori antigens for vaccine development, the amino

CC acid sequences predicted from various ORF were analysed for significant

CC homology to other known or exported membrane proteins. Having identified

CC and determined the sequences of interest, particular regions can be

CC isolated from H. pylori by PCR amplification for recombinant polypeptide

XX production, e.g. in E. coli hosts.

XX SQ Sequence 407 AA;

Query Match

Best Local Similarity 5.5%; Score 71; DB 18; Length 407;

Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 SHAAFTTIVIPAIVGGTATGVTGSLLSWGLKQAEANKTPDKPKVWRIQAGKF 90  
|||||  
Db 47 shaaffttviipaivggatgvtgslgllswgllkqaeanktpdkpkvriqagkf 106  
|||||

OY 91 NEFPNKEYDLY 101  
|||||

Db 107 nefpnkeydly 117  
|||||

## RESULT 12

AAW20321  
ID AAW20321 standard; protein; 90 AA.

XX AC

AAW20321;

DT XX 09-JUL-1997 (first entry)

DE XX H. pylori protein.

XX KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;  
XX OS Helicobacter pylori.

XX PN WO9640893-A1.

XX PD 19-DEC-1996.

XX PF 06-JUN-1996; 96WO-US09122.

XX PR 01-APR-1996; 96US-0630405.

XX PR 07-JUN-1995; 95US-0487032.

XX PA (ASTR ) ASTRA AB.

XX PI Berglindh OT, Smith D, Mellgaard BL;

XX DR WPI; 1997-052306/05.

XX DR N-PSDB; AAT67525.

XX XX Helicobacter pylori nucleic acid sequences and related

PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PT infection, and to detect Helicobacter

XX PS Disclosure; Pages 514; 1481pp; English.

XX CC The present sequence is a Helicobacter pylori protein of unknown

CC function. The protein may be used in a vaccine to prevent or treat

CC H. pylori infection or to identify H. pylori polypeptide binding

CC compounds, useful as potential H. pylori life cycle activators or

CC inhibitors. The genomic sequence of H. pylori (ATCC 55679) was

CC determined from overlapping contigs generated by mechanically shearing

CC the bacterial DNA. The sequences were analysed for ORF of at least 180

CC nucleotides, and the predicted coding regions defined by computer

CC evaluation. To identify likely H. pylori antigens for vaccine

CC development, the amino acid sequences predicted from various ORF were

CC analysed for significant homology to other known or exported membrane

CC proteins. Having identified and determined the sequences of interest,

CC particular regions can be isolated from H. pylori by PCR amplification

CC for recombinant polypeptide production, e.g. in E. coli hosts.

XX SQ Sequence 90 AA;

Query Match 4.6%; Score 59; DB 18; Length 90;

Best Local Similarity 100.0%; Pred. No. 5.6e-51;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 SHAAFTTIVIPAIVGGTATGVTGSLLSWGLKQAEANKTPDKPKVWRIQAGK 89  
|||||  
Db 31 shaaffttviipaivggatgvtgslgllswgllkqaeanktpdkpkvriqagk 89  
|||||

## RESULT 13

AAW89963  
ID AAW89963 standard; protein; 216 AA.

XX AC AAW89963;

XX AC AAW89963;

DT 18-FEB-1999 (first entry)

XX DE Antigen from cluster 62.

XX XX

KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;  
 KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma.  
 XX Helicobacter pylori.

OS W09849314-A2.

PN 05-NOV-1998.

XX 27-APR-1998; 98WO-US08487.

XX 14-OCT-1997; 97US-0061958.

PR 25-APR-1997; 97US-0045107.

XX (GENE-) GENELABS TECHNOLOGIES INC.  
 PA Chow TP, Fry KE, Lim MY, McAtee CP;

PI WPI; 1999-009433/01.

XX New Helicobacter pylori antigens and related nucleic acid sequences  
 XX - useful in serological diagnosis and protective vaccines, providing  
 PT long-lasting immune response

PT Claim 1: Page 272-273; 402pp; English.

XX The present sequence represents a Helicobacter pylori antigenic protein  
 CC that is characterised by immunoreactivity with H. pylori-positive  
 CC antisera. The proteins are highly immunogenic and induce a long-lasting  
 CC immune response that persists even after antimicrobial treatment. In  
 CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are  
 CC highly sensitive and specific. The specification also describes 69  
 CC previously unrecognised immunogenic cluster families. H. pylori antigens  
 CC are used to detect H. pylori-specific antibodies, for diagnosing  
 CC infection or to confirm eradication of infection, and in vaccines to  
 CC protect against H. pylori infection and related diseases (gastritis,  
 CC peptic ulcer, gastric adenocarcinoma/lymphoma).

XX Sequence 216 AA;

Query Match 4.6%; Score 59; DB 20; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-50;  
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 797 QSMVNPDPNFKYLICKAWKNIGISKTAGSKISVYVYLGNSPTENGNTNLTPTNTTSN 855

DB 30 qsmvnpdpnykyligkawkngisktagkskisyvylgnsptengntnltptnttsn 88

RESULT 14

AAW27720

ID AAW27720 standard; Protein; 288 AA.

XX AAW27720;

DT 11-MAY-1998 (first entry)

XX H. pylori VacA protein autotransporter region.

DE VacA gene; autotransporter; diagnostic; therapy;

KW Gram-negative bacteria; surface presented polypeptide.

XX Helicobacter pylori.

OS Key Location/Qualifiers

XX Protein 1..288

FT /note= "partial protein sequence"

XX W09735022-A1.

XX 25-SEP-1997.

XX

PF 15-MAR-1996; 96WO-EP01130.

XX 15-MAR-1996; 96WO-EP01130.

XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX Jose J, Maurer J, Meyer TF;

PI WPI; 1997-480227/44.

DR N-PSDB; AAT88157.

XX Presentation of peptide(s) on surface of Gram-negative bacteria -  
 PT via transformation with vector encoding signal peptide, presented  
 PT peptide and transporter domain of auto-transporter, producing  
 PT peptide libraries for epitope mapping

XX Claim 8; Fig 24; 84pp; German.

XX This sequence represents an autotransporter membrane integration  
 CC region from the H. pylori VacA gene. This region is involved in a novel  
 CC method which allows the presentation of stable fusion polypeptides on the  
 CC surface of Gram-negative bacteria which can be released into the  
 CC surrounding media. The method can be used to produce a variegated  
 CC population of surface-presented polypeptides, so that bacteria expressing  
 CC polypeptides with particular properties can be identified and  
 CC simultaneously selected, e.g. for epitope mapping or selection of ligands  
 CC with the highest affinity for antibodies, major histocompatibility  
 CC complex (MHC) molecules or other components of the immune system.  
 CC Selected polypeptides can be used diagnostically, e.g. to screen sera or  
 CC antibody banks, and polypeptide expressing cells may be used as live  
 CC vaccines. They may be used therapeutically, e.g. when the polypeptide is  
 CC an antibody, to remove or concentrate pollutants, inactivate toxins,  
 CC prepare and process food, prepare washing compositions and label cells.  
 CC Selected bacteria can be stored, reproduced and replicated on a large  
 CC scale as individual clones.

XX Sequence 288 AA;

Query Match 3.7%; Score 48; DB 18; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-39;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1082 ANSLNSGANNTFVYSRIFANQHEHDFEAQAGLGSQSSLNFKSALL 1129

DB 74 anslnsganntfvyrsrifanqhehdfefagaglgsgsslnfksall 121

RESULT 15

AAB52506

ID AAB52506 standard; Protein; 224 AA.

XX AAB52506;

DT 23-FEB-2001 (first entry)

XX Helicobacter pylori bait polypeptide #24.

DE Helicobacter pylori; two-hybrid system; protein-protein interaction;

KW bait polypeptide; gastric ulcer; antibacterial.

XX Helicobacter pylori.

OS W0200066722-A1.

XX 09-NOV-2000.

XX 14-APR-2000; 2000WO-IB00603.

XX 30-APR-1999; 99EP-0401066.

XX (HYBR-) HYBRIGENICS SA.

XX



PI Legrain P, Sellig L, Rain J;  
 XX WPI; 2000-687535/67.  
 XX N-PSDB; AAC97252.  
 XX  
 PT A two-hybrid system for identifying compounds useful in the treatment  
 PT of e.g. gastric ulcers comprises producing a collection of recombinant  
 PT cell clones -  
 XX  
 PS Example 5; Page 106-107; 267pp; English.  
 XX  
 CC The present sequence is a bait polypeptide used in a Helicobacter  
 CC pylori two-hybrid screen to identify protein-protein interactions.  
 CC The method is used to identify a recombinant cell clone expressing a  
 CC prey polypeptide which is capable of interacting with the bait  
 CC polypeptide. The two hybrid system is useful for screening compounds  
 CC for antibacterial activity. It may be used in the treatment of gastric  
 CC ulcers. The polynucleotides are useful as amplification primers or  
 CC specific detection probes. The polypeptides, vectors or host cells can  
 CC be used as immunogens to produce mono- or polyclonal antibodies. The  
 CC polynucleotides, polypeptides, antibodies, vectors, host cells or  
 CC modulating agents can be used to produce a pharmaceutical composition.  
 XX  
 SQ Sequence 224 AA;

Query Match 3.5%; Score 45; DB 21; Length 224;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-36;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVKLMGVNMGRLQYVGAYLAPSYSTINTSKVTGEVNFHNLTVGD 294  
 |||||  
 DB 86 svklmgvnmgrlqyvgaylapytintskvtgevnfhnlvtgd 130

RESULT 16  
 AAW20260  
 ID AAW20260 standard; Protein; 141 AA.  
 AC AAW20260;  
 XX  
 DT 30-JUL-1997 (first entry).  
 XX  
 DE H. pylori protein.  
 XX  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
 XX  
 OS Helicobacter pylori.  
 PN WO9640893-A1.  
 PD 19-DEC-1996.  
 XX  
 PF 06-JUN-1996; 96WO-US09122.  
 XX  
 PR 01-APR-1996; 96US-0630405.  
 PR 07-JUN-1995; 95US-0487032.  
 XX  
 PA (ASTR ) ASTRA AB.  
 PI Berglindh OT, Smith D, Mellgaard BL;  
 XX WPI; 1997-052306/05.  
 DR N-PSDB; AAT67759.  
 XX  
 PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter  
 XX  
 PS Disclosure; Page 463-464; 1481pp; English.  
 XX

CC This sequence is a H. pylori protein.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.  
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 XX  
 SQ Sequence 141 AA;

Query Match 2.9%; Score 37; DB 18; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-28;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1146 RASYGYDFAFFRNALVKPSGVSYNHLGSTNFKSNS 1182  
 |||||  
 DB 51 rasygydaffrnalvlpkpsvgsynhlgstnfksns 87

RESULT 17  
 AAW24621  
 ID AAW24621 standard; Protein; 141 AA.  
 XX  
 AC AAW24621;  
 XX  
 DT 11-AUG-1997 (first entry)  
 XX  
 DE H. pylori protein.  
 XX  
 KW Transmembrane; cytoplasmic; cell envelope; flagella; transport;  
 KW secreted; periplasmic; chronic gastritis; duodenal ulcer disease;  
 KW activator; inhibitor; bacterial life cycle; vaccine; immunise;  
 KW detection; antisense; inhibition.  
 XX  
 OS Helicobacter pylori.  
 PN WO9719098-A1.  
 XX  
 PD 29-MAY-1997.  
 XX  
 PF 15-NOV-1996; 96WO-US18542.  
 PR 17-NOV-1995; 95US-0561469.  
 XX  
 PA (ASTR ) ASTRA AB.  
 PI Smith DH;  
 XX  
 DR WPI; 1997-298052/27.  
 DR N-PSDB; AAT77439.  
 XX  
 PT Helicobacter pylori nucleic acid sequences and related proteins -  
 PT used for diagnostics and therapeutics  
 XX  
 PS Disclosure; Page 152-153; 1481pp; English.  
 XX  
 CC This sequence is a H. pylori protein.  
 CC Helicobacter pylori has been strongly linked to chronic gastritis and  
 CC duodenal ulcer disease. The nucleic acid sequences of the invention  
 CC are used to evaluate compounds, especially activators or inhibitors of  
 CC bacterial life cycle, for the ability to bind an H. pylori nucleic acid  
 CC sequence. The nucleic acid sequences, and corresponding proteins, are  
 CC also useful for generating vaccines for immunising subjects against H.  
 CC pylori or for use in detecting the presence of Helicobacter species in  
 CC a sample. Antisense nucleic acid sequences of these sequences are

used to inhibit expression of a gene from *Helicobacter* species. H. pylori whole genomic DNA was isolated and nebulised to a median size of 2000 bp. Purified DNA fragments were blunt-ended and ligated to unique BstXI-linker adapters in 100-1000 fold molar excess. These linkers are complementary to the BstXI-cut pMPX vectors, while the overhang is not self-complementary. Therefore the linkers will not concatamerise nor will the cut vector re-ligate itself easily. The linker-adaptor inserts were ligated to each of the 20 pMPX vectors to construct a series of shotgun subclone libraries. The purified DNA samples were then sequenced.

Note: The ORF/protein reference number for this sequence was obtained from the related specification, W09640893.

XX SQ Sequence 141 AA;

Query Match 2.9%; Score 37; DB 18; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.1e-28;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1146 RASGYDFAFRNALVKPSGVSYNHLGSTNFKNS 1182  
|||||  
Db 51 rasygydaffrnalvkpsgvsynhlgstnfksns 87

## RESULT 18

AAW55324  
ID AAW55324 standard; Protein; 211 AA.

XX AC AAW55324;

XX DT 15-JUN-1998 (first entry)

XX DE H. pylori ORF 07apl1015orf4 protein.

XX KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
KW identification; binding compound; bacteria; life cycle; activator;  
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.

XX OS *Helicobacter pylori*.

XX PN W09737044-A1.

XX PD 09-OCT-1997.

XX PF 27-MAR-1997; 97WO-US05223.

XX PR 06-DEC-1996; 96US-0761318.

XX PR 29-MAR-1996; 96US-0625811.

XX PR 02-APR-1996; 96US-0758731.

XX PR 25-OCT-1996; 96US-0736905.

XX PR 28-OCT-1996; 96US-0738859.

XX PA (ASTR ) ASTRA-AB.

XX PI Alm RA, Smith D;

XX DR WPI; 1997-503122/46.

XX DR N-PSDB; AAV24733.

XX PT *Helicobacter pylori* nucleic acid sequences and encoded  
XX polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
XX infection and for diagnosis of H. pylori infection

XX PS Claim 14; Pages 550-551; 1145pp; English.

XX This sequence is a H. pylori protein of unspecified function.  
XX The protein may be used in a vaccine to prevent or treat H. pylori  
XX infection or to identify H. pylori polypeptide binding compounds,  
XX useful as potential H. pylori life cycle activators or inhibitors. The  
XX DNA and probes derived from it may be used for the identification of  
XX H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
XX acid sequences complementary to the DNA act as antisense sequences and

CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
CC against the protein can be used in immunoassays to evaluate the abundance  
CC and distribution of H. pylori-specific antigens. The genomic sequence of  
CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
CC by mechanically shearing the bacterial DNA. The sequences were analysed  
CC for ORF of at least 180 nucleotides, and the predicted coding regions  
CC defined by computer evaluation. To identify likely H. pylori antigens for  
CC vaccine development, the amino acid sequences predicted from various ORF  
CC were analysed for significant homology to other known or exported  
CC membrane proteins. Having identified and determined the sequences of  
CC interest, particular regions can be isolated from H. pylori by PCR  
CC amplification for recombinant polypeptide production, e.g. in *E. coli*  
XX hosts.

SQ Sequence 211 AA;

Query Match 2.9%; Score 37; DB 18; Length 211;  
Best Local Similarity 100.0%; Pred. No. 1.6e-28;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1146 RASGYDFAFRNALVKPSGVSYNHLGSTNFKNS 1182  
|||||  
Db 121 rasygydaffrnalvkpsgvsynhlgstnfksns 157

## RESULT 19

AAW20755

XX ID AAW20755 standard; Protein; 211 AA.

XX AC AAW20755;

XX DT 15-JUL-1997 (first entry)

XX DE H. pylori protein.

XX KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;  
KW outer membrane; cell envelope; transporter.

XX OS *Helicobacter pylori*.

XX PN W09640893-A1.

XX PD 19-DEC-1996.

XX PF 06-JUN-1996; 96WO-US09122.

XX PR 01-APR-1996; 96US-0630405.

XX PR 07-JUN-1995; 95US-0487032.

XX PA (ASTR ) ASTRA-AB.

XX PI Berglindh OT, Smith D, Mellgaard BL;

XX DR WPI; 1997-052306/05.

XX DR N-PSDB; AAT68008.

XX PT *Helicobacter pylori* nucleic acid sequences and related  
XX polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
XX infection, and to detect *Helicobacter*

XX PS Disclosure; Pages 1168-1169; 1481pp; English.

XX The present sequence is a *Helicobacter pylori* protein of unknown  
XX function. The protein may be used in a vaccine to prevent or treat  
XX H. pylori infection or to identify H. pylori polypeptide binding  
XX compounds, useful as potential H. pylori life cycle activators or  
XX inhibitors. The genomic sequence of H. pylori (ATCC 55679) was  
XX determined from overlapping contigs generated by mechanically shearing  
XX the bacterial DNA. The sequences were analysed for ORF of at least 180  
XX nucleotides, and the predicted coding regions defined by computer

CC evaluation. To identify likely H. pylori antigens for vaccine  
 CC development, the amino acid sequences predicted from various ORF were  
 CC analysed for significant homology to other known or exported membrane  
 CC proteins. Having identified and determined the sequences of interest,  
 CC particular regions can be isolated from H. pylori by PCR amplification  
 CC for recombinant polypeptide production, e.g. in E. coli hosts.

XX SQ Sequence 211 AA;

Query Match 2.9%; Score 37; DB 18; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-28;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1146 RASGYDFAFRNALVLKPSGVSYNHLGSTNFKNS 1182  
 |||||  
 Db 121 rasygydfafrrnalvlkpsgvsvynhlgstnfksns 157

RESULT 20  
 AAW55462  
 ID AAW55462 standard; Protein; 257 AA.  
 XX AAW55462;  
 AC AAW55462;  
 XX 24-JUN-1998 (first entry)  
 XX H. pylori ORF 07apl1015\_23938312\_f3\_2 cell envelope OMP.  
 XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.  
 XX Helicobacter pylori.  
 XX W09737044-AL.  
 XX 09-OCT-1997.  
 XX 27-MAR-1997; 97WO-US05223.  
 XX 06-DEC-1996; 96US-0761318.  
 PR 29-MAR-1996; 96US-0625811.  
 PR 02-APR-1996; 96US-0758731.  
 PR 25-OCT-1996; 96US-0736905.  
 PR 28-OCT-1996; 96US-0738859.  
 XX (ASTR ) ASTRA AB.  
 XX Alm RA, Smith D;  
 XX WPI; 1997-503122/46.  
 DR N-PSDB; AAV24871.  
 XX Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX Disclosure; Page 667-668; 1145pp; English.

CC This sequence is a H. pylori cell envelope outer membrane protein  
 CC (OMP) having a terminal Phe residue.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed

CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.

XX SQ Sequence 257 AA;

Query Match 2.9%; Score 37; DB 18; Length 257;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-28;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1146 RASGYDFAFRNALVLKPSGVSYNHLGSTNFKNS 1182  
 |||||  
 Db 107 rasygydfafrrnalvlkpsgvsvynhlgstnfksns 143

RESULT 21  
 AAW55437  
 ID AAW55437 standard; Protein; 133 AA.  
 XX AAW55437;  
 AC AAW55437;  
 XX 22-JUN-1998 (first entry)  
 XX H. pylori ORF hp5p15580orf1 protein.  
 DE Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.  
 XX Helicobacter pylori.  
 XX W09737044-AL.  
 XX 09-OCT-1997.  
 XX 27-MAR-1997; 97WO-US05223.  
 XX 06-DEC-1996; 96US-0761318.  
 PR 29-MAR-1996; 96US-0625811.  
 PR 02-APR-1996; 96US-0758731.  
 PR 25-OCT-1996; 96US-0736905.  
 PR 28-OCT-1996; 96US-0738859.  
 XX (ASTR ) ASTRA AB.  
 XX Alm RA, Smith D;  
 XX WPI; 1997-503122/46.  
 DR N-PSDB; AAV24846.  
 XX Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX Claim 14; Page 645-646; 1145pp; English.

CC This sequence is a H. pylori protein of unspecified function.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated

CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.

XX Sequence 133 AA;

Query Match 2.6%; Score 34; DB 18; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 1e-25;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 561 KYNIGSVGYTHFSEIDIGSQSRINTVRLTGTR 594  
 Db 3 kngisvgythfseidigsqsrintvrltgtrs 36

RESULT 22

AAW20382  
 ID AAW20382 standard; protein; 63 AA.

XX AC AAW20382;

XX DT 14-JUL-1997 (first entry)

XX DE H. pylori derived protein of unspecified function.

XX KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.

XX OS Helicobacter pylori.

XX FH Key Location/Qualifiers

FT Misc-difference 43  
 FT /label= Unknown  
 FT /note= "encoded by AGY"

FT Misc-difference 57  
 FT /label= Unknown

FT Misc-difference 63  
 FT /note= "encoded by YCA"

FT /note= "encoded by GMG"

XX PN WO9640893-A1.

XX PD 19-DEC-1996.

XX PF 06-JUN-1996; 96WO-US09122.

XX PR 01-APR-1996; 96US-0630405.

XX PR 07-JUN-1995; 95US-0487032.

XX PA (ASTR ) ASTRA AB.

XX PI Berglinth OT, Smith D, Mellgaard BL;

XX DR WPI; 1997-052306/05.

XX DR N-PSDB; AAT67569.

XX PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter

XX PS Disclosure; Page 568; 1481pp; English.

XX CC The present sequence is a H. pylori derived protein, no further details  
 CC are given in the specification.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori.

CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.  
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.

XX Sequence 63 AA;

Query Match 2.2%; Score 29; DB 18; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-21;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 201 VEINRVSGAGRKASSTVLTQASEGIT 229

Db 14 veinnrvsgagrkasstvtlqasegit 42

RESULT 23

AAAR41353

ID AAR41353 standard; Protein; 23 AA.

XX AC AAR41353;

XX DT 04-MAR-1994 (first entry)

XX DE N-terminal sequence of vacuolating toxin.

XX KW Helicobacter pylori; antigen; antigenic composition; vacuolation;  
 KW toxin; peptic ulcer disease; gastric malignancy; vaccine; immunity.

XX OS Helicobacter pylori.

XX PN WO9316723-A.

XX PD 02-SEP-1993.

XX PF 24-FEB-1993; 93WO-US01558.

XX PR 26-FEB-1992; 92US-0841644.

XX PA (UYVA-) UNIV VANDERBILT.

XX PI Blaser MJ, Cover TL;

XX DR WPI; 1993-288125/36.

XX PT Pure CB antigen from Helicobacter pylori - used to protect  
 PT against infection and for detection to determine susceptibility  
 PT to peptic ulcer or gastric malignancy

XX PS Claim 3; Page 49-50; 76pp; English.

XX CC An antigen (CB antigen) contains this N-terminal peptide and can be  
 CC used for inducing protective immunity against H. pylori infection  
 CC and for the detection and production of antibodies. The antigen, of  
 CC antibodies and nucleic acids can be used to detect the presence of  
 CC infection with H. pylori and thereby identify patients at risk of  
 CC developing peptic ulcer disease or gastric malignancy.

XX Sequence 23 AA;

Query Match 1.8%; Score 23; DB 14; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-15;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AFTTVPVPAIVGGIATGTAVGT 56  
 Db 1 affttvpaivggiatgtavgt 23

## RESULT 24

AAW20362  
 ID AAW20362 standard; Protein; 93 AA.

XX AC AAW20362;  
 XX DT 14-JUL-1997 (first entry)  
 XX DE H. pylori protein.  
 XX KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
 XX OS Helicobacter pylori.

XX PN WO9640893-A1.

XX PD 19-DEC-1996.

XX PF 06-JUN-1996; 96WO-US09122.

XX PR 01-APR-1996; 96US-0630405.

XX PR 07-JUN-1995; 95US-0487032.

XX PA (ASTR ) ASTRA AB.

XX PI Berglindh OT, Smith D, Mellgaard BL;

XX DR WPI; 1997-052306/05.

XX DR N-PSDB; AAT67354.

XX PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter  
 XX Disclosure; Page 551; 1481pp; English.

XX CC This sequence is a H. pylori protein.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.  
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.

XX SQ Sequence 93 AA;

Query Match 1.5%; Score 20; DB 18; Length 93;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1263 LAKEVFLNLGVVYLNHLISN 1282

Db 60 lakevflnlgvvylhnlisn 79

## RESULT 25

AAW55322

ID AAW55322 standard; Protein; 95 AA.

XX AC AAW55322;

XX DT 15-JUN-1998 (first entry)

XX DE H. pylori ORF 07apl1015orf2 protein.

XX KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.

XX OS Helicobacter pylori.

XX PN WO9737044-A1.

XX PD 09-OCT-1997.

XX PF 27-MAR-1997; 97WO-US05223.

XX PR 06-DEC-1996; 96US-0761318.

XX PR 29-MAR-1996; 96US-0625811.

XX PR 02-APR-1996; 96US-0758731.

XX PR 25-OCT-1996; 96US-0736905.

XX PR 28-OCT-1996; 96US-0738859.

XX PA (ASTR ) ASTRA AB.

XX PI Alm RA, Smith D;

XX DR WPI; 1997-503122/46.

XX DR N-PSDB; AAV24731.

XX PT Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX Claim 14; Page 549; 1145pp; English.

XX CC This sequence is a H. pylori protein of unspecified function.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.

XX SQ Sequence 95 AA;

Query Match 1.5%; Score 20; DB 18; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1263 LAKEVFLNLGVVYLNHLISN 1282

Db 62 lakevflnlgvvylhnlisn 81

## RESULT 26



PF 24-FEB-1993; 93WO-US01558.  
 XX  
 PR 26-FEB-1992; 92US-0841644.  
 XX  
 PA (UVA-) UNIV VANDERBILT.  
 XX  
 PI Blaser MJ, Cover TL;  
 XX  
 DR WPI; 1993-288125/36.  
 XX  
 PT Pure CB antigen from Helicobacter pylori - used to protect  
 PT against infection and for detection to determine susceptibility  
 PT to peptic ulcer or gastric malignancy  
 XX  
 PS Disclosure; Page 28; 76pp; English.  
 XX  
 CC The CB antigen from Helicobacter pylori was purified by column  
 CC chromatography and then digested by Arg-c protease. The resulting  
 CC peptides were then separated on a Vydac C18 column. These peptides  
 CC were sequenced. Some of the results were used to generate degenerate  
 CC oligonucleotide primers (AA046150-52) which were then used to amplify  
 CC sequences of the CB antigen (vacuolating toxin gene).  
 XX  
 SQ Sequence 16 AA;  
 Query Match 1.0%; Score 13; DB 14; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-05; Mismatches 0; Indels 0; Gaps 0;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 168 GPFNGNSFTSYKD 180  
 DB 1 gqfngnsftsykd 13  
 RESULT 29  
 AAW20215  
 ID AAW20215 standard; Protein; 15 AA.  
 XX  
 AC AAW20215;  
 XX  
 DT 09-JUL-1997 (first entry)  
 XX  
 DE H. pylori derived protein.  
 XX  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacterium; life cycle; activator;  
 KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;  
 KW diagnosis.  
 XX  
 OS Helicobacter pylori.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 13  
 FT /label= Unknown  
 FT /note= "encoded by YTC"  
 XX  
 PN W09640893-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 06-JUN-1996; 96WO-US09122.  
 XX  
 PR 01-APR-1996; 96US-0630405.  
 PR 07-JUN-1995; 95US-0487032.  
 XX  
 PA (ASTR ) ASTRA AB.  
 XX  
 PI Berglindh OT, Smith D, Mellgaard BL;  
 XX  
 DR WPI; 1997-052306/05.  
 DR N-PSDB; AAT67441.  
 XX

PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter  
 XX  
 PS Disclosure; Page 426; 1481pp; English.  
 XX  
 CC This sequence represents a H. pylori protein of unknown function.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.  
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 140 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 XX  
 SQ Sequence 15 AA;  
 Query Match 0.8%; Score 10; DB 18; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.017; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 134 VDMKDAVGTY 143  
 DB 1 vdmkdavgtly 10  
 RESULT 30  
 AAW20825  
 ID AAW20825 standard; Protein; 246 AA.  
 XX  
 AC AAW20825;  
 XX  
 DT 16-JUL-1997 (first entry)  
 XX  
 DE H. pylori cytoplasmic protein 11cpl2006orf17.  
 XX  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN W09640893-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 06-JUN-1996; 96WO-US09122.  
 XX  
 PR 01-APR-1996; 96US-0630405.  
 PR 07-JUN-1995; 95US-0487032.  
 XX  
 PA (ASTR ) ASTRA AB.  
 XX  
 PI Berglindh OT, Smith D, Mellgaard BL;  
 XX  
 DR WPI; 1997-052306/05.  
 DR N-PSDB; AAW20825.  
 XX  
 PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter  
 XX  
 PS Claim 61; Page 1230; 1481pp; English.  
 XX  
 CC The present sequence shows a Helicobacter pylori cytoplasmic protein  
 CC that may be used in a vaccine to prevent or treat H. pylori

CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.  
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.

XX Sequence 246 AA;

Query Match 0.6%; Score 8; DB 18; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
 DB 185 nrvgsgag 192  
 |||||

RESULT 31  
 AAY11003  
 ID AAY11003 standard; Protein; 246 AA.  
 XX AC AAY11003;  
 XX DT 08-JUN-1999 (first entry)  
 DE H. pylori ORF hp2e10911\_960952\_c2\_86 cellular protein.  
 XX Vaccine; probe; diagnostic; ORF; cell envelope protein;  
 KW secreted protein; cellular protein.  
 KW Helicobacter pylori.

XX WO9818323-Al.  
 PN 07-MAY-1998.  
 PD 28-OCT-1997; 97WO-US19575.  
 PF 14-JUL-1997; 97US-0891928.  
 PR 28-OCT-1996; 96US-0739150.  
 PR 06-DEC-1996; 96US-0759739.

XX (ASTR ) ASTRA AB.

XX Alm RA, Smith D;  
 FI WPI; 1998-271811/24.  
 DR N-PSDB; AAX30470.

XX Helicobacter pylori nucleic acids and proteins - used to develop  
 PT products for the detection, prevention and treatment of H. pylori  
 PT infections

XX Claims 27, 31; Page 213-214; 279pp; English.

CC Recombinant or substantially pure preparations of H. pylori polypeptides  
 CC are disclosed, together with the nucleic acids encoding them. In all,  
 CC 73 ORFs are shown. The proteins are variously cell envelope proteins,  
 CC secreted proteins or other cellular proteins. Vaccines containing the  
 CC nucleic acids or proteins are claimed, as are probes containing at least  
 CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful  
 CC for treating or reducing the risk of H. pylori infections, and the  
 CC probes can be used diagnostically for detecting the presence of  
 CC Helicobacter in a sample. The products are also of use in screening  
 CC for compounds having the ability to interfere with the H. pylori life

CC cycle or to inhibit H. pylori infection.  
 XX Sequence 246 AA;

Query Match 0.6%; Score 8; DB 19; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
 DB 185 nrvgsgag 192  
 |||||

RESULT 32  
 AAW80647  
 ID AAW80647 standard; Protein; 293 AA.

XX AC AAW80647;  
 XX DT 24-DEC-1998 (first entry)  
 XX DE S. pneumoniae transport protein.

XX Streptococcus pneumoniae protein; recombinant; gene expression; DNA chip;  
 KW virulence; antibody; infection; detection; treatment; hypothetical;  
 KW cell wall biosynthetic, external target; minimal gene set protein.

XX Streptococcus pneumoniae.

XX WO9826072-Al.

XX 18-JUN-1998.

XX 09-DEC-1997; 97WO-US252578.

XX 13-DEC-1996; 96US-0036281.

XX (ELIL ) LILLY & CO ELI.

XX Baltz RH, Burgett SG, Dehoff BS, Hoskins JA, Jaskunas SR;  
 PI Mills BJ, Norris FH, Peery RB, Rostock PK, Rostock PR;  
 PI Skatrud PL, Smith MC, Solenberg PJ, Treadway PJ;  
 PI Young Bellido ML;

XX WPI; 1998-348529/30.  
 DR N-PSDB; AAV65221.

XX Streptococcus pneumoniae nucleic acid sequences - used in DNA chips  
 PT for evaluating gene expression, and identification of virulence  
 PT genes

XX Claim 3; Pages 234-235; 333pp; English.

XX This sequence represents a Streptococcus pneumoniae transport  
 CC protein. The invention provides DNA sequences (AAV65201 to AAV65304)  
 CC from the Streptococcus pneumoniae genome and corresponding protein  
 CC sequences (AAW80605 to AAW80728). The protein sequences are classified as  
 CC hypothetical, cell wall biosynthetic, external target, or minimal gene  
 CC set proteins. A recombinant host containing a vector comprising any of  
 CC the above nucleic acids can be used for the recombinant expression of the  
 CC proteins. The invention also provides a DNA chip having arrayed on it at  
 CC least 15 base pair fragment of any one or more of these DNA sequences.  
 CC The DNA chip can be used for evaluating gene expression in S.  
 CC pneumoniae and for identifying virulence genes in S. pneumoniae.  
 CC Antibodies that selectively bind to the above proteins or peptide  
 CC fragments can be used to treat S. pneumoniae infection. The antibodies  
 CC can also be used to detect S. pneumoniae cells.

XX Sequence 293 AA;

Query Match 0.6%; Score 8; DB 19; Length 293;



Best Local Similarity 100.0%; Pred. No. 28;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1127 ALLQDLNQ 1134  
|||||||  
Db 189 allqdlng 196

## RESULT 33

AAW20372  
ID AAW20372 standard; protein; 325 AA.

XX AC AAW20372;

XX 29-JUL-1997 (first entry)

DE H. pylori cytoplasmic protein, 291700.aa.

XX Vaccine; prevention; treatment; infection; identification;  
KW binding compound; bacterium; life cycle; activator; bacteria;  
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;  
KW cytoplasmic.

OS Helicobacter pylori.

XX Key Location/Qualifiers

FT Misc-difference 5 /label= unknown

FT /note= "encoded by CYA"

FT Misc-difference 9 /label= unknown

FT /note= "encoded by ATR"

FT Misc-difference 17 /label= unknown

FT /note= "encoded by GRA"

FT Misc-difference 211 /label= unknown

FT /note= "encoded by AGK"

FT Misc-difference 234 /note= "encoded by ACS"

FT Misc-difference 240 /note= "encoded by ACS"

FT /note= "encoded by ACS"

FT Misc-difference 244 /label= unknown

FT /note= "encoded by AAW"

FT Misc-difference 247 /note= "encoded by GGS"

FT Misc-difference 253 /note= "encoded by GAR"

XX WO9640893-A1.

PN 19-DEC-1996.

XX 06-JUN-1996; 96WO-US09122.

XX 01-APR-1996; 96US-0630405.

PR 07-JUN-1995; 95US-0487032.

XX (ASTR ) ASTRA AB.

PA Berglindh OT, Smith D, Mellgaard BL;

PI WPI; 1997-052306/05.

XX N-PSDB; AAT67782.

DR Helicobacter pylori nucleic acid sequences and related

XX polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PT infection, and to detect Helicobacter

XX Claim 61; Pages 558-559; 1481pp; English.

PS The present sequence is a Helicobacter pylori cytoplasmic protein.

CC

CC The protein may be used in a vaccine to prevent or treat  
CC H. pylori infection or to identify H. pylori polypeptide binding  
CC compounds, useful as potential H. pylori life cycle activators or  
CC inhibitors. The genomic sequence of H. pylori (ATCC 55679) was  
CC determined from overlapping contigs generated by mechanically  
CC shearing the bacterial DNA. The sequences were analysed for ORF of  
CC at least 180 nucleotides, and the predicted coding regions defined  
CC by computer evaluation. To identify likely H. pylori antigens for  
CC vaccine development, the amino acid sequences predicted from  
CC various ORF were analysed for significant homology to other known  
CC or exported membrane proteins. Having identified and determined  
CC the sequences of interest, particular regions can be isolated from  
CC H. pylori by PCR amplification for recombinant polypeptide  
CC production, e.g. in E. coli hosts.

XX SQ Sequence 325 AA;

Query Match 0.6%; Score 8; DB 18; Length 325;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212

Db 153 nrvgsgag 160

## RESULT 34

AAW24644  
ID AAW24644 standard; Protein; 325 AA.

XX AC AAW24644;

XX 11-AUG-1997 (first entry)

DE H. pylori cytoplasmic protein, 291700.aa.

XX Chronic gastritis; duodenal ulcer disease; activator;

XX inhibitor; bacterial life cycle; vaccine; immunisation; detection;

XX antisense; inhibition; cytoplasmic; vaca.

XX Helicobacter pylori.

XX Key Location/Qualifiers

FT Misc-difference 5 /note= "encoded by CYA"

FT Misc-difference 9 /note= "encoded by ATR"

FT Misc-difference 17 /note= "encoded by GRA"

FT Misc-difference 211 /note= "encoded by ACK"

FT Misc-difference 244 /note= "encoded by AAW"

XX WO9719098-A1.

PN 29-MAY-1997.

XX 15-NOV-1996; 96WO-US18542.

XX 17-NOV-1995; 95US-0561469.

XX (ASTR ) ASTRA AB.

PA Smith DH;

PI WPI; 1997-298052/27.

XX N-PSDB; AAT77462.

DR Helicobacter pylori nucleic acid sequences and related proteins -

XX used for diagnostics and therapeutics

PT

XX

PS Claim 18; Pages 166-167; 235pp; English.

XX The present sequence is a Helicobacter pylori cytoplasmic protein, which was found to be homologous to vacA following BLAST protein analysis.

CC H. pylori has been strongly linked to chronic gastritis and duodenal ulcer disease. The nucleic acid sequences of the invention are used to evaluate compounds, especially activators or inhibitors of bacterial life cycle, for the ability to bind an H. pylori nucleic acid sequence. The nucleic acid sequences, and corresponding proteins, are also useful for generating vaccines for immunising subjects against H. pylori or for use in detecting the presence of Helicobacter species in a sample. Antisense nucleic acid sequences of these sequences are used to inhibit expression of a gene from Helicobacter species. H. pylori whole genomic DNA was isolated and nebulised to a median size of 2000 bp. Purified DNA fragments were blunt-ended and ligated to unique BstXI-linker adapters in 100-1000 fold molar excess. These linkers are complementary to the BstXI-cut pMPX vectors, while the overhang is not self-complementary. Therefore the linkers will not concatamerise nor will the cut vector re-ligate itself easily. The linker-adaptor inserts were ligated to each of the 20 pMPX vectors to construct a series of shotgun subclone libraries. The purified DNA samples were then sequenced.

CC Note: The ORF/protein reference number for this sequence was obtained from the related specification, WO9640893.

XX Sequence 325 AA;

SQ

Query Match 0.6%; Score 8; DB 18; Length 325;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
|||||||

Db 153 nrvgsgag 160  
|||||||

RESULT 35  
AAW55735  
ID AAW55735 standard; Protein; 1213 AA.  
XX  
AC AAW55735;  
XX  
DT 13-JUL-1998 (first entry)  
XX  
DE H. pylori ORF 07ee50709\_960952\_f2\_47 secreted protein.  
XX  
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
KW identification; binding compound; bacteria; life cycle; activator;  
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;  
XX secreted protein.  
OS Helicobacter pylori.  
XX  
PN W09737044-A1.  
XX  
PD 09-OCT-1997.  
XX  
PF 27-MAR-1997; 97WO-US05223.  
XX  
PR 06-DEC-1996; 96US-0761318.  
PR 29-MAR-1996; 96US-0625811.  
PR 02-APR-1996; 96US-0758731.  
PR 25-OCT-1996; 96US-0736905.  
PR 28-OCT-1996; 96US-0738859.  
XX  
XX (ASTR ) ASTRA AB.  
PA  
XX Alm RA, Smith D;  
PI  
XX WPI; 1997-503122/46.  
XX

DR N-PSDB; AAV25144.

XX Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) - useful in vaccines to treat or prevent H. pylori infection and for diagnosis of H. pylori infection

PT  
PT  
XX Disclosure; Page 1007-1010; 1145pp; English.

PS  
XX This sequence represents a Helicobacter pylori secreted protein. The protein may be used in a vaccine to prevent or treat H. pylori infection or to identify H. pylori polypeptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors. The DNA and probes derived from it may be used for the identification of H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic acid sequences complementary to the DNA act as antisense sequences and can be used to prevent the translation of H. pylori mRNA. Antibodies against the protein can be used in immunoassays to evaluate the abundance and distribution of H. pylori-specific antigens. The genomic sequence of H. pylori (ATCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.

XX  
SQ Sequence 1213 AA;

Query Match 0.6%; Score 8; DB 18; Length 1213;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
|||||||

Db 185 nrvgsgag 192  
|||||||

RESULT 36  
AAU00023  
ID AAU00023 standard; Protein; 1851 AA.  
XX  
AC AAU00023;  
XX  
DT 10-MAY-2001 (first entry)  
XX  
DE Human activated T-lymphocyte associated sequence 2, ATLAS-2.  
XX  
KW Human; activated T-lymphocyte associated sequence 2; ATLAS-2; antibody;  
KW cytokine receptor; autoimmune disorder; immune disorder; cancer;  
KW T-lymphocyte-associated disorder; cell proliferation disorder; tumour;  
KW cell differentiation disorder; immune deficiency disorder; malignancy;  
KW viral infection; bacterial infection; fungal infection; metabolism;  
KW chromosome 11p15.5.  
XX  
OS Homo sapiens.  
XX  
PN WO200114564-A2.  
XX  
PD 01-MAR-2001.  
XX  
PF 18-AUG-2000; 2000WO-US22699.  
XX  
PR 20-AUG-1999; 99US-0150105.  
PR 28-APR-2000; 2000US-0560101.  
PR 28-APR-2000; 2000US-0560365.  
PR 28-APR-2000; 2000US-0560948.  
PR 28-APR-2000; 2000US-0561533.  
XX  
XX (CURA-) CURAGEN CORP.  
PA

PA (BIOJ ) BIOGEN INC.  
 PI Peyman JA, Green CD, Hsu A, Browning JA, Carulli J;  
 XX WPI; 2001-218453/22.  
 DR N-PSDB; AAS00033.  
 XX  
 XX New isolated activated T lymphocyte associated sequences for treating  
 PT or preventing immune system associated disorders such as autoimmune  
 PT disorder, immune disorder, and T-lymphocyte-associated disorder  
 XX  
 XX Claim 14; Fig 2; 114pp; English.  
 XX  
 CC The sequence represents human activated T-lymphocyte associated sequence  
 CC 2, ATLAS-2. ATLAS-2 is related by homology to cytokine receptors and its  
 CC gene is located on chromosome 11p15.5. ATLAS proteins, polynucleotides  
 CC and antibodies are useful for treating/preventing conditions associated  
 CC with an autoimmune disorder, immune disorder, T-lymphocyte-associated  
 CC disorder, cell-proliferation disorder, cell differentiation disorder,  
 CC and immune deficiency disorder and for screening for a modulator of  
 CC activity or of latency or predisposition to an immune disorder.  
 CC ATLAS proteins, polynucleotides and antibodies are useful in therapeutic  
 CC or prophylactic treatment of diseases associated with cell proliferation  
 CC (e.g. cancers, malignancies and tumours). The polynucleotides are useful  
 CC in gene therapy, to detect ATLAS mRNA or a genetic lesion in an ATLAS  
 CC gene, to modulate ATLAS activity, to screen drugs or compounds that  
 CC modulate ATLAS activity or expression and to treat disorders  
 CC characterised by insufficient or excessive production of ATLAS protein or  
 CC production of ATLAS protein forms that have decreased or aberrant  
 CC activity compared to ATLAS wild type protein and in tissue typing to  
 CC identify individuals. The antibodies are useful for localisation/  
 CC quantitation, isolation and detection of ATLAS and to monitor protein  
 CC levels in tissue. ATLAS is useful for treating/preventing infection by  
 CC bacteria, viruses and fungi, affecting bodily characteristics, e.g.  
 CC biorhythms, fertility or metabolism, affecting behavioural  
 CC characteristics, and for providing analgesic effects. A host cell  
 CC containing the polynucleotide is useful to produce non-human transgenic  
 CC animals.  
 XX  
 SQ Sequence 1851 AA;  
 Query Match 0.6%; Score 8; DB 22; Length 1851;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 965 GLQTLISLS 972  
 Db 897 GLQTLISLS 904  
 RESULT 37  
 AAW98828  
 ID AAW98828 standard; Protein; 2893 AA.  
 XX  
 AC AAW98828;  
 XX  
 DT 31-MAR-1999 (first entry)  
 XX  
 DE H. pylori GHPO 1484 protein.  
 XX  
 DE GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9843478-A1.  
 XX  
 PD 08-OCT-1998.  
 XX  
 PF 01-APR-1998; 98WO-US05371.  
 XX  
 PF 29-JUL-1997; 97US-0902615.  
 XX

PR 01-APR-1997; 97US-0833457.  
 PR 24-JUN-1997; 97US-0881227.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 XX  
 PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
 XX WPI; 1998-542293/46.  
 DR N-PSDB; AAX14547.  
 XX  
 PT New isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 PT infections and gastrointestinal diseases  
 XX  
 PS Claim 8; Page 1827-1840; 2054pp; English.  
 CC  
 CC This sequence represents a Helicobacter pylori GHPO protein of the  
 CC invention. The polypeptides can be used for preventing or treating  
 CC Helicobacter infections, and gastroduodenal diseases associated with  
 CC these infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
 CC used for the production of antibodies. The products can also be used for  
 CC detection and diagnosis.  
 XX  
 SQ Sequence 2893 AA;  
 Query Match 0.6%; Score 8; DB 19; Length 2893;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 205 NRVGSGAG 212  
 Db 176 nrvgsag 183  
 RESULT 38  
 AAW71556  
 ID AAW71556 standard; Protein; 2893 AA.  
 XX  
 AC AAW71556;  
 XX  
 DT 09-NOV-1998 (first entry)  
 XX  
 DE Helicobacter polypeptide GHPO 1484.  
 XX  
 KW GHPO 1484; infection; therapy; diagnosis; vaccine; gastritis;  
 KW ulcer.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9821225-A1.  
 XX  
 PD 22-MAY-1998.  
 XX  
 PF 14-NOV-1997; 97WO-US21353.  
 XX  
 PR 29-JUL-1997; 97US-0902615.  
 PR 14-NOV-1996; 96US-0749051.  
 PR 01-APR-1997; 97US-0831309.  
 PR 01-APR-1997; 97US-0833457.  
 PR 01-APR-1997; 97US-0834705.  
 PR 24-JUN-1997; 97US-0881227.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 XX  
 PI Al-Garawi A, Haas R, Kleanthous H, Meyer T, Miller C;  
 PI Odenbreit S, Tomb J;  
 XX WPI; 1998-297855/26.  
 DR

DR N-PSDB; AAV52091.  
 XX  
 PT Helicobacter polynucleotide and polypeptide sequences - useful to  
 PT treat or prevent gastrointestinal infection  
 XX  
 PS  
 PS Claim 1; Page 330-337; 362pp; English.  
 XX  
 CC This claimed Helicobacter pylori polypeptide, designated GHPO 1484,  
 CC can be used in vaccination methods for preventing or treating  
 CC Helicobacter infection. 85 Helicobacter polypeptides (see  
 CC AAW1474-W71558) are claimed, as well as isolated polynucleotides  
 CC (see AAV52009-93) that encode them. The invention also provides:  
 CC methods for producing these Helicobacter polypeptides in  
 CC recombinant host systems, and related expression cassettes, vectors  
 CC and transformed or transfected host cells; live vaccine vectors  
 CC that contain the polynucleotides of the invention and which can be  
 CC used to prevent or treat Helicobacter infection; therapeutic and/or  
 CC prophylactic methods involving administration of polynucleotide  
 CC molecules, polypeptides or monospecific antibodies; methods for  
 CC detecting the presence of Helicobacter in samples using e.g.  
 CC the polypeptides or monospecific antibodies; and methods for  
 CC purifying the polypeptides by antibody-based affinity  
 CC chromatography.  
 XX  
 SQ Sequence 2893 AA;

Query Match 0.6%; Score 8; DB 19; Length 2893;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
 Db 176 nrvgsgag 183  
 |||||

RESULT 39  
 AAB46351  
 ID AAB46351 standard; Protein; 2902 AA.  
 XX  
 AC AAB46351;  
 XX  
 DT 05-APR-2001 (first entry)  
 XX  
 DE H. pylori HPN165 protein.  
 XX  
 KW Microbial infection; antibacterial; Helicobacter pylori infection;  
 KW vaccine; screening.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO200073502-A2.  
 PD 07-DEC-2000.  
 XX  
 PF 31-MAY-2000; 2000WO-EP05024.  
 XX  
 PR 31-MAY-1999; 99DE-1024965.  
 PR 17-JUN-1999; 99DE-1027740.  
 PR 21-JUL-1999; 99DE-1034029.  
 XX  
 XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 XX (CREA-) CREATOGEN GMBH.  
 XX  
 PI Apfel H, Fuchs TM, Gibbs CP, Hueck CJ, Meyer TF;  
 XX  
 DR WPI; 2001-049948/06.  
 DR N-PSDB; AAF25628.  
 XX  
 XX Preparing an agent for diagnosis or control of microbial infection,  
 PT useful particularly against Helicobacter, based on identification of  
 PT essential genes in defective mutants -

PS Claim 37; Figure 15; 366pp; German.

XX  
 CC This invention describes a novel preparation of an agent (A) for  
 CC detection, prevention and/or treatment of microbial infection by:  
 CC (i) identifying essential genes (I) and corresponding polypeptides  
 CC (II); (ii) identifying compounds that are directed against (II) and  
 CC inactivate the microbe; (iii) testing these for suitability for use; and  
 CC (iv) formulating selected (A). Identifying essential genes (I) comprises  
 CC preparation of gene-deficient microorganisms by conditional antisense  
 CC inhibition (CAI) and/or subtractive recombination mutagenesis (SRM),  
 CC then determining viability and/or survival of the deficient organisms.  
 CC The products of the invention have antibacterial activity. (A) which may  
 CC be a nucleic acid (Ia), vector or host cell containing (Ia), derived  
 CC polypeptide (IIa), or fragments, (IIa)-specific antibodies or their  
 CC fragments or an inhibitor of (IIa) are particularly used for diagnosis,  
 CC treatment or prevention of infection by Helicobacter pylori. Particularly  
 CC (Ia) and (IIa) are used in DNA, subunit or live vaccines. The method  
 CC identifies essential genes, including those that have homologs in other  
 CC species, so identified (A) should have a broad spectrum of activity. Many  
 CC gene-deficient cells can be screened quickly, in an automated process,  
 CC and the identified genes can be used for screening without purification.  
 XX  
 SQ Sequence 2902 AA;

Query Match 0.6%; Score 8; DB 22; Length 2902;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
 Db 185 nrvgsgag 192  
 |||||

RESULT 40  
 AAR41351  
 ID AAR41351 standard; Protein; 8 AA.  
 XX  
 AC AAR41351;  
 XX  
 DT 04-MAR-1994 (first entry)  
 XX  
 DE Peptide fragment of CB antigen.  
 XX  
 KW Helicobacter pylori; antigen; antigenic composition; vacuolation;  
 KW toxin; peptic ulcer disease; gastric malignancy; vaccine; immunity;  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9316723-A.  
 PD 02-SEP-1993.  
 XX  
 PF 24-FEB-1993; 93WO-US01558.  
 XX  
 PR 26-FEB-1992; 92US-0841644.  
 XX  
 XX (UYVA-) UNIV VANDERBILT.  
 PA  
 XX Blaser MJ, Cover TL;  
 XX  
 DR WPI; 1993-288125/36.  
 XX  
 PT Pure CB antigen from Helicobacter pylori - used to protect  
 PT against infection and for detection to determine susceptibility  
 PT to peptic ulcer or gastric malignancy  
 XX  
 PS Disclosure; Page 28; 76pp; English.

XX The CB antigen from Helicobacter pylori was purified by column  
 CC chromatography and then digested by Arg-c protease. The resulting  
 CC peptides were then separated on a Vydac C18 column. These peptides  
 CC were sequenced. Some of the results were used to generate degenerate

CC oligonucleotide primers (AA046150-52) which were then used to amplify  
 CC sequences of the CB antigen (vacuolating toxin gene).  
 XX  
 SQ Sequence 8 AA;

Query Match 0.5%; Score 7; DB 14; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 IKNVEIT 632  
 |||||  
 Db 1 iknveit 7

## RESULT 41

AAR84029  
 ID AAR84029 standard; peptide; 13 AA.

XX  
 AC AAR84029;

XX 14-MAY-1996 (first entry)

XX Murine MHC class II binding peptide I-E(alpha) 54-66.

XX Murine; MHC class II; binding peptide; haptenated peptides;  
 KW contact; sensitivity; desensitising; mammal; allergen; ivy;  
 KW urushiol; poison; oak.

OS Synthetic.

PN WO9526980-A2.

XX 12-OCT-1995.

XX 30-MAR-1995; 95WO-US04121.

XX 06-FEB-1995; 95US-0383645.

XX 01-APR-1994; 94US-0222206.

XX (IMMU-) IMMULOGIC PHARM CORP.

PI Gefter ML, Gelber C, Greenstein JL, Hackett CJ;  
 PI Wilson KJ;

XX WPI; 1995-358583/46.

XX Haptenated peptide(s) capable of binding to Class II MHC molecules -  
 PT for treating contact dermatitis

PS Example; Fig 2; 85pp; English.

XX A peptide of 7-30 amino acids capable of binding to a murine MHC  
 CC class II mol. (i.e. AAR84018-47) covalently linked to 1-3 hapten  
 CC mols. can be used for treating contact sensitivity, or  
 CC desensitising a mammal to a contact allergen (e.g. urushiol of  
 CC poison ivy/oak). The peptide-hapten cpds. disrupt the normal  
 CC proliferation of hapten-specific T cells, or alter the T cell  
 CC mediated delayed-type hypersensitivity response to the hapten,  
 CC resulting in effective desensitisation to the hapten.

XX Sequence 13 AA;

Query Match 0.5%; Score 7; DB 16; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQCAL 1115  
 |||||  
 Db 1 feaqqal 7

## RESULT 42

AAR84032  
 ID AAR84032 standard; peptide; 13 AA.

XX  
 AC AAR84032;

XX 14-MAY-1996 (first entry)

XX Murine MHC class II binding peptide I-E(alpha) 54-66, K63.

XX Murine; MHC class II; binding peptide; haptenated peptides;  
 KW contact; sensitivity; desensitising; mammal; allergen; ivy;  
 KW urushiol; poison; oak.

OS Synthetic.

XX Key Location/Qualifiers  
 FT Modified-site 1

FT Modified-site 13 /note= "acylated"

FT Modified-site 13 /note= "amidated"

XX WO9526980-A2.

XX 12-OCT-1995.

XX 30-MAR-1995; 95WO-US04121.

XX 06-FEB-1995; 95US-0383645.

XX 01-APR-1994; 94US-0222206.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Gefter ML, Gelber C, Greenstein JL, Hackett CJ;  
 PI Wilson KJ;

XX WPI; 1995-358583/46.

XX Haptenated peptide(s) capable of binding to Class II MHC molecules -  
 PT for treating contact dermatitis

PS Example; Fig 2; 85pp; English.

XX A peptide of 7-30 amino acids capable of binding to a murine MHC  
 CC class II mol. (i.e. AAR84018-47) covalently linked to 1-3 hapten  
 CC mols. can be used for treating contact sensitivity, or  
 CC desensitising a mammal to a contact allergen (e.g. urushiol of  
 CC poison ivy/oak). The peptide-hapten cpds. disrupt the normal  
 CC proliferation of hapten-specific T cells, or alter the T cell  
 CC mediated delayed-type hypersensitivity response to the hapten,  
 CC resulting in effective desensitisation to the hapten.

XX Sequence 13 AA;

Query Match 0.5%; Score 7; DB 16; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQCAL 1115  
 |||||  
 Db 1 feaqqal 7

## RESULT 43

AAR84035  
 ID AAR84035 standard; peptide; 16 AA.

XX  
 AC AAR84035;

XX 14-MAY-1996 (first entry)

XX Murine MHC class II binding peptide I-E(alpha) 54-66, K63 (PDC).

XX Murine; MHC class II; binding peptide; haptenated peptides;  
 KW contact; sensitivity; desensitising; mammal; allergen; ivy;  
 KW urushiol; poison; oak.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "acylated"

FT Modified-site 16 /note= "amidated"

FT FT

XX W09526980-A2.

XX 12-OCT-1995.

XX 30-MAR-1995; 95WO-US04121.

XX 06-FEB-1995; 95US-0383645.

XX 01-APR-1994; 94US-0222206.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Gefter ML, Gelber C, Greenstein JL, Hackett CJ;  
 PI Wilson KJ;

XX WPI; 1995-358583/46.

XX Haptenated peptide(s) capable of binding to Class II MHC molecules -  
 PT for treating contact dermatitis

XX Example; Fig 2; 85pp; English.

XX A peptide of 7-30 amino acids capable of binding to a murine MHC  
 CC class II mol. (i.e. AAR84018-47) covalently linked to 1-3 hapten  
 CC mols. can be used for treating contact sensitivity, or  
 CC desensitising a mammal to a contact allergen (e.g. urushiol of  
 CC poison ivy/oak). The peptide-hapten cpds, disrupt the normal  
 CC proliferation of hapten-specific T cells, or alter the T cell  
 CC mediated delayed-type hypersensitivity response to the hapten,  
 CC resulting in effective desensitisation to the hapten.

XX Sequence 16 AA;

Query Match 0.5%; Score 7; DB 16; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 FEAQGA 1115

Db 1 feaqaal 7

RESULT 44

AAW87824  
 ID AAW87824 standard; Peptide; 16 AA.

XX AAW87824;

XX 10-MAR-1999 (first entry)

XX Epitope of a Bcl-2 associated protein designated Bax.

XX Bcl-2 associated protein; Bax; bcl-2; antibody; modulator;  
 KW bcl-2-related function; apoptosis.

XX Synthetic.

OS Homo sapiens.

OS Mus sp.

XX US5856171-A.

XX 05-JAN-1999.  
 PD 10-NOV-1994; 94US-0337646.  
 XX 10-NOV-1994; 94US-0337646.  
 PR 26-AUG-1993; 93US-0112208.  
 PR 25-MAY-1994; 94US-0248819.  
 XX (UNIW ) UNIV WASHINGTON.

XX Korsmeyer SJ;

XX WPI; 1999-105119/09.

XX DNA composition encoding bcl-2 two-hybrid and reporter system - for  
 PT identifying modulators of bcl-2 function

XX Disclosure; Column 33; 105pp; English.

XX AAW87814-31 represent epitopes derived from a Bcl-2 associated protein  
 CC designated Bax. The peptides are used to raise antibodies, which  
 CC are used in a composition of the invention. The composition  
 CC comprises a bcl-2 family member polypeptide, a naturally occurring  
 CC Bax polypeptide and an antibody that binds to the Bax polypeptide.  
 CC The composition is used to identify modulators of bcl-2-related  
 CC function, e.g. substances that inhibit binding of Bax to bcl-2,  
 CC which would be potentially useful as drugs for modulating  
 CC apoptosis.

XX Sequence 16 AA;

Query Match 0.5%; Score 7; DB 20; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22

Db 2 slalvga 8

RESULT 45

AAW874141

ID AAW874141 standard; Peptide; 16 AA.

XX AAW874141;

XX 22-MAY-2001 (first entry)

XX Bax epitope #11.

XX Bax; cytostatic; immunosuppressive; immunostimulant; infection;  
 KW apoptosis modulator; bcl-2 associated x protein; cancer therapy; AIDS;  
 KW autoimmunity; immunodeficiency; reperfusion injury; stroke; aging;  
 KW myocardial infarction; traumatic brain injury; ischaemia; Bcl-2;  
 KW neurodegenerative diseases; hepatitis; transplant rejection; toxemia;  
 KW lymphoproliferative disease.

XX Unidentified.

XX US6184202-B1.

XX 06-FEB-2001.

XX 11-SEP-1997; 97US-0927326.

XX 10-NOV-1994; 94US-0337646.

XX 26-AUG-1993; 93US-0112208.

XX 25-MAY-1994; 94US-0248819.

XX (UNIW ) UNIV WASHINGTON.

PI Korsmeyer SJ;  
 DR WPI; 2001-256104/26.  
 XX  
 XX  
 PT Modulating apoptosis of a cell, useful in maintaining homeostasis in  
 PT adult tissues, or treating proliferative or autoimmune diseases,  
 PT comprises administering a bcl-2 polypeptide that interacts with a 21 kD  
 PT bcl-2 associated x protein -  
 XX  
 XX Disclosure: Columns 33-34; 105pp; English.  
 XX  
 XX The present invention relates to a method of modulating apoptosis of a  
 CC cell. The method comprises administering to the cell an agent,  
 CC comprising a Bhl domain or BH2 domain, capable of modulating formation of  
 CC at least one complex selected from bcl-2:bcl-2 complexes, bcl-Xl:bcl-Xl  
 CC complexes, bcl-2 associated x protein (Bax):Bax complexes, bcl-2:Bax  
 CC complexes or bcl-Xl:Bax complexes. Modulating apoptosis is especially  
 CC useful in cancer therapy, and treating autoimmunity, immunodeficiency  
 CC diseases (e.g. AIDS), reperfusion injury, myocardial infarction, stroke,  
 CC traumatic brain injury, neurodegenerative diseases, aging, ischaemia,  
 CC toxemia, infection, hepatitis, transplant rejection, and  
 CC lymphoproliferative diseases. The present sequence is a peptide, which  
 CC was used in the method of the present invention.  
 XX  
 XX Sequence 16 AA;

Query Match 0.5%; Score 7; DB 22; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 |||||  
 Db 2 slalvga 8

RESULT 46  
 AAR82531  
 ID AAR82531 standard; peptide; 17 AA.  
 XX  
 AC AAR82531;  
 XX  
 DT 15-APR-1996 (first entry)  
 XX  
 DE MHC groove specific peptide #3.  
 XX  
 KW Major histocompatibility complex; MHC; T-cell receptor; TCR;  
 KW autoimmune disease; immunodeficiency disease; immune response;  
 KW immunoproliferation disease; graft-host rejection; therapy.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 12  
 FT /label- Iso  
 FT  
 XX WO9523814-A1.  
 PN  
 XX 08-SEP-1995.  
 PD  
 XX 03-MAR-1995; 95WO-US02689.  
 PF  
 XX 04-MAR-1994; 94US-0207481.  
 PR  
 XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.  
 PA  
 XX Kappler JW, Marrack P;  
 PI WPI; 1995-320543/41.  
 XX  
 XX Peptide-MHC complex comprising antigenic peptide, linker and MHC  
 PT segment - useful as reagents for the treatment of diseases including  
 PT auto-immune diseases, immuno-stimulatory diseases or graft-host

PT rejection  
 XX  
 PS Disclosure: Page 18; 94pp; English.  
 XX  
 XX The sequences represented by AAR82527, AAR82528 and AAR82531 are  
 CC antigenic peptides that bind specifically to the groove of major  
 CC histocompatibility complexes (MHC). The grooves of a MHC protein are  
 CC internal peptide binding sites. Once the peptide has bound to the MHC  
 CC protein, the resulting peptide-MHC complex can bind to a T-cell receptor  
 CC (TCR). These peptides bind to MHC proteins that are involved in  
 CC autoimmune diseases, immunodeficiency diseases, immunoproliferation  
 CC diseases, and graft-host rejection. These peptides (and other peptides  
 CC that stimulate antigenic responses) can be joined to an MHC molecule by a  
 CC linker. These complexes may be used to regulate an immune response. The  
 CC complexes are capable of being recognised by a TCR alone or in  
 CC combination with additional MHC proteins. These complexes are useful for  
 CC therapeutic purposes and experimental purposes. They can also be used as  
 CC reagents for the treatment of diseases including autoimmune diseases,  
 CC immunodeficiency diseases, immunoproliferation diseases, and graft-host  
 CC rejection.  
 XX  
 SQ Sequence 17 AA;

Query Match 0.5%; Score 7; DB 16; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 PEAQCAL 1115  
 |||||  
 Db 3 feagga 9

RESULT 47  
 AAW34191  
 ID AAW34191 standard; peptide; 17 AA.  
 XX  
 AC AAW34191;  
 XX  
 DT 06-MAY-1998 (first entry)  
 XX  
 DE Bt-E.alpha.  
 XX  
 KW MHC class II molecule; major histocompatibility complex; T cell receptor;  
 KW myelin basic protein; MBP; modified antigenic peptide; myasthenia gravis;  
 KW rheumatoid arthritis; therapy; autoimmune disease; allograft rejection;  
 KW systemic lupus erythematosus; insulin-dependent diabetes; cell typing;  
 KW multiple sclerosis; allergy; autoreactive T cell.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "biotinylated"  
 FT Modified-site 17  
 FT /note= "C-terminal amide"  
 FT  
 XX WO9740852-A1.  
 PN  
 XX 06-NOV-1997.  
 PD  
 XX 18-MAR-1997; 97WO-US04360.  
 PF  
 XX 30-APR-1996; 96US-0640344.  
 PR  
 XX (ANER-) ANERGEN INC.  
 PA  
 XX Deshpande S, Mukku P, Nag B;  
 PI WPI; 1997-549492/50.  
 XX  
 XX Modified peptide antigen for major histocompatibility molecule -  
 PT used specifically to inactivate T cell receptors implicated in

PT auto:immune disease  
XX  
PS Example; Page 25; 45pp; English.  
XX  
CC This sequence represents a modified fragment of the E.alpha protein.  
CC This sequence can be used in the modified antigenic peptide of the  
CC invention. The modified antigenic peptide (A) is for a major  
CC histocompatibility complex (MHC) class II molecule, and has the amino  
CC acid structure Y-Z, where Y = a hydrophobic amino acid (aa), or a series  
CC of 1-5 aa of which at least one is hydrophobic; and Z = a peptide epitope  
CC for the MHC class II molecule. (A) has a higher affinity for the MHC  
CC class II molecule than the peptide epitope. When a complex containing (A)  
CC can bind to, and inactivate, a T cell receptor it is useful for  
CC preventing or treating disorders associated with autoimmune diseases,  
CC particularly autoimmune diseases, specifically rheumatoid arthritis (RA),  
CC systemic lupus erythematosus (SLE), insulin-dependent diabetes,  
CC myasthenia gravis (MG) and multiple sclerosis (MS). The complex  
CC containing (A) can also be used for preventing or treating allergy and  
CC allograft rejection. The complex can be used to detect specific  
CC autoreactive T cells, from their ability to bind to the complex. The  
CC complex can also be used for in vivo/in vitro diagnosis, including T cell  
CC typing, isolation or labelling of specific cells, to assay potential  
CC inhibitors of MHC cell interactions, for imaging and monitoring  
CC treatment. Addition of Y improves bonding, probably by providing an  
CC anchor within the MHC antigen-binding pocket, and thus provides higher  
CC occupancy of this pocket.  
XX  
SQ Sequence 17 AA;

Query Match 0.5%; Score 7; DB 18; Length 17;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
Db 3 feaqal 9  
|||||

RESULT 48  
AAW34192  
ID AAW34192 standard; peptide; 17 AA.  
AC AAW34192;  
XX  
DT 06-MAY-1998 (first entry)  
DE Bt-E.alpha+Y.  
XX  
KW MHC class II molecule; major histocompatibility complex; T cell receptor;  
KW myelin basic protein; MBP; modified antigenic peptide; myasthenia gravis;  
KW rheumatoid arthritis; therapy; autoimmune disease; allograft rejection;  
KW systemic lupus erythematosus; insulin-dependent diabetes; cell typing;  
KW multiple sclerosis; allergy; autoreactive T cell.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT Modified-site 17 /note= "biotinylated"  
FT Modified-site 17 /note= "C-terminal amide"  
XX  
PN W09740852-A1.  
XX  
PD 06-NOV-1997.  
XX  
PF 18-MAR-1997; 97WO-US04360.  
XX  
PR 30-APR-1996; 96US-0640344.  
XX  
PA (ANER-) ANERGEN INC.  
XX  
PI

PI Deshpande S, Mukku P, Nag B;  
XX WPI; 1997-549492/50.  
DR  
XX  
PT Modified peptide antigen for major histocompatibility molecule -  
PT used specifically to inactivate T cell receptors implicated in  
PT auto:immune disease  
XX  
PS Example; Page 25; 45pp; English.  
XX  
CC This sequence represents a modified fragment of the E.alpha protein.  
CC This sequence can be used in the modified antigenic peptide of the  
CC invention. The modified antigenic peptide (A) is for a major  
CC histocompatibility complex (MHC) class II molecule, and has the amino  
CC acid structure Y-Z, where Y = a hydrophobic amino acid (aa), or a series  
CC of 1-5 aa of which at least one is hydrophobic; and Z = a peptide epitope  
CC for the MHC class II molecule. (A) has a higher affinity for the MHC  
CC class II molecule than the peptide epitope. When a complex containing (A)  
CC can bind to, and inactivate, a T cell receptor it is useful for  
CC preventing or treating disorders associated with autoreactive T cells,  
CC particularly autoimmune diseases, specifically rheumatoid arthritis (RA),  
CC systemic lupus erythematosus (SLE), insulin-dependent diabetes,  
CC myasthenia gravis (MG) and multiple sclerosis (MS). The complex  
CC containing (A) can also be used for preventing or treating allergy and  
CC allograft rejection. The complex can be used to detect specific  
CC autoreactive T cells, from their ability to bind to the complex. The  
CC complex can also be used for in vivo/in vitro diagnosis, including T cell  
CC typing, isolation or labelling of specific cells, to assay potential  
CC inhibitors of MHC cell interactions, for imaging and monitoring  
CC treatment. Addition of Y improves bonding, probably by providing an  
CC anchor within the MHC antigen-binding pocket, and thus provides higher  
CC occupancy of this pocket.  
XX  
SQ Sequence 17 AA;

Query Match 0.5%; Score 7; DB 18; Length 17;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
Db 3 feaqal 9  
|||||

RESULT 49  
AAW79534  
ID AAY79534 standard; Peptide; 17 AA.  
XX  
AC AAY79534;  
XX  
DT 15-AUG-2000 (first entry)  
XX  
DE Mouse I-E molecule thrid hyper V region self-peptide I alpha 52.  
XX  
KW Antigen presenting cell; T-lymphocyte; T-cell; immunomodulator;  
KW autoimmune disease; allergy; cancer; infection; graft rejection;  
KW immunotherapy; therapy; mouse; major histocompatibility complex;  
KW MHC.  
XX  
OS Mus musculus.  
XX  
PN W0200023053-A2.  
XX  
PD 27-APR-2000.  
XX  
PF 19-OCT-1999; 99WO-US24666.  
XX  
PR 20-OCT-1998; 98US-0105018.  
XX  
PA (ALBA/) ALBANI S.  
XX  
PI Albani S;



XX WPI; 2000-339492/29.  
 XX  
 XX New artificial antigen presenting cells useful for isolating and  
 PT expanding T cells, and modulating T cell responses for the treatment of  
 PT e.g. autoimmune diseases, allergies  
 XX  
 XX Example 2; Page 62; 179pp; English.  
 PS  
 XX  
 XX Self-peptide I alpha 52 corresponds to residues 52-68 (third hyper  
 CC V region) of the alpha chain of the mouse I-E molecule. It was  
 CC used in an experiment to investigate the effects of a naturally  
 CC processed self-peptide on the maintenance and proliferation of T  
 CC cells which may cross-react with homologous peptides of exogenous  
 CC origin. The experiment helped to define the antigen specificity of  
 CC a positively selected T cell population in a non-transgenic BALB/c  
 CC model. The results emphasized the fact that without a method  
 CC such as that of the current invention to capture T cells, it is not  
 CC possible to evaluate polyclonal antigen specific T cell selection  
 CC a non-transgenic system. The present peptide was also used in  
 CC experiments to identify cross-reactive T cells with specificity  
 CC for homologous peptides, and to identify T cell receptor usage by  
 CC antigen specific T cells. The invention is directed to artificial  
 CC antigen presenting cells (APC) and methods of making APC. These  
 CC are used to isolate and expand T cell populations and to modulate  
 CC T cell responses. The invention also provides novel methods for  
 CC the identification and isolation and antigen-specific T cells. The  
 CC methods provide for the construction of liposomes containing  
 CC MHC:peptide complexes, accessory molecules, co-stimulatory  
 CC molecules, adhesion molecules, and other molecules irrelevant to T  
 CC cell binding or modulation that are used in the binding of  
 CC artificial APC to solid support systems that may be used in the  
 CC retrieval and identification and antigen-specific T cells. Devices  
 CC and methods are provided for treating conditions that would benefit  
 CC from modulation of T cell response, e.g. autoimmune disorders  
 CC (especially type I diabetes mellitus, multiple sclerosis,  
 CC rheumatoid arthritis, dermatomyositis, juvenile rheumatoid  
 CC arthritis and uveitis), allergies, cancer, viral infections, and  
 CC graft rejection.  
 XX  
 XX Sequence 17 AA;  
 SQ  
 Query Match 0.5%; Score 7; DB 21; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1109 FEAQAL 1115  
 Db 3 feaqal 9  
 RESULT 50  
 AAB48952  
 ID AAB48952 standard; Protein; 17 AA.  
 XX  
 XX AAB48952;  
 AC  
 XX 27-MAR-2001 (first entry)  
 DT  
 XX I-Ab-restricted MHC class II epitope, SEQ ID NO:8.  
 DE  
 XX Transposable element; MHC epitope; major histocompatibility complex;  
 KW intracellular bacterial pathogen; loxP site; Cre recombinase;  
 KW insertion end; in-frame fusion; detection; antigen;  
 KW disseminated insertions of class-II epitopes; DICE-II; transposon Tn5;  
 KW I-Ab-restricted MHC class II epitope.  
 XX  
 XX Unidentified.  
 OS  
 XX  
 XX WO200071158-A1.  
 PN  
 XX 30-NOV-2000.  
 PD

XX  
 XX 26-MAY-2000; 2000WO-US14587.  
 XX  
 XX 26-MAY-1999; 99US-0136210.  
 PR  
 XX (UYOR-) UNIV OREGON HEALTH SCI.  
 PA  
 XX Heffron FL, Parker DC, Ellefson DD;  
 PI  
 XX WPI; 2001-031967/04.  
 DR  
 XX Transposable element for detecting an antigenic epitope of a pathogen,  
 PT comprising 5' and 3' recombining sites, nucleic acid sequences encoding  
 PT a selectable marker and major histocompatibility complex (MHC) epitope,  
 PT and an insertion end  
 XX  
 XX Claim 5; Page 41; 63pp; English.  
 PS  
 XX The invention relates to a novel transposable element comprising DNA  
 CC encoding a selectable marker (e.g., antibiotic resistance) located  
 CC between a 5' recombining site and a 3' recombining site (e.g., loxP  
 CC sites); DNA encoding an MHC (major histocompatibility complex) epitope  
 CC either 5' of the 5' recombining site or 3' of the 3' recombining site;  
 CC and insertion ends comprising an inverted repeat sequence at the 5' and  
 CC 3' ends of the transposable element sufficient for integration of the  
 CC transposable element. The transposable elements of the invention are able  
 CC to introduce in-frame insertions throughout the chromosome of an  
 CC intracellular bacterial pathogen. This system "tags" the bacterial gene  
 CC and resulting protein, allowing the identification of proteins  
 CC secreted across the membranes of the eukaryotic cell infected by the  
 CC bacterium. In one embodiment, the transposable elements contain an  
 CC antibiotic resistance cassette, two minimal loxP recombination sites, an  
 CC MHC class I or class II epitope, and flanking insertion ends. A  
 CC transposase, such as the Cre recombinase protein, is expressed in trans  
 CC from a plasmid, or can be included in the transposable element. The Cre  
 CC recombinase loops out the intervening sequences containing the antibiotic  
 CC resistance cassette. When the transposable element inserts within a gene,  
 CC the resolved insertion places the MHC class I or class II epitope in  
 CC frame with the gene. The transposable elements of the invention are  
 CC useful for detecting an antigenic epitope of an intracellular bacterial  
 CC pathogen, such as *Salmonella* sp., *Mycobacterium tuberculosis* and *Listeria*  
 CC monocytes. Certain embodiments of the technology, termed  
 CC "disseminated insertions of class-I epitopes" (DICE-I; DICE-II for  
 CC class II epitopes) allow the rapid and accurate identification of  
 CC proteins involved in bacterial pathogenesis so that such proteins can  
 CC be used as vaccine and drug targets. Carrier vaccines may be generated  
 CC by infecting bacteria with a transposable element of the invention  
 CC which additionally comprises an antigen associated with a disease,  
 CC preferably cancer or a viral or bacterial disease, operably linked to the  
 CC MHC epitope DNA of the transposable element. The present sequence  
 CC represents an I-Ab-restricted MHC class II epitope specifically claimed  
 CC for use in the invention.  
 XX  
 XX Sequence 17 AA;  
 SQ  
 Query Match 0.5%; Score 7; DB 22; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1109 FEAQAL 1115  
 Db 3 feaqal 9  
 RESULT 51  
 AAR41352  
 ID AAR41352 standard; Protein; 20 AA.  
 XX  
 XX AAR41352;  
 AC  
 XX 04-MAR-1994 (first entry)  
 DT  
 XX

DE Peptide fragment of CB antigen.  
 XX Helicobacter pylori; antigen; antigenic composition; vacuolation;  
 KW toxin; peptic ulcer disease; gastric malignancy; vaccine; immunity;  
 XX Helicobacter pylori.  
 OS Helicobacter pylori.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2 /note= "Valine or Isoleucine."  
 FT  
 XX  
 XX WO9316723-A.  
 XX  
 PD 02-SEP-1993.  
 XX  
 XX 24-FEB-1993; 93WO-US01558.  
 XX  
 PR 26-FEB-1992; 92US-0841644.  
 XX  
 PA (UYVA-) UNIV VANDERBILT.  
 XX  
 PI Blaser MJ, Cover TL;  
 XX  
 DR WPI; 1993-288125/36.  
 XX  
 XX Pure CB antigen from Helicobacter pylori - used to protect  
 PT against infection and for detection to determine susceptibility  
 PT to peptic ulcer or gastric malignancy  
 XX  
 PS Disclosure; Page 28; 76pp; English.  
 XX  
 CC The CB antigen from Helicobacter pylori was purified by column  
 CC chromatography and then digested by Arg-c protease. The resulting  
 CC peptides were then separated on a Vydac C18 column. These peptides  
 CC were sequenced. Some of the results were used to generate degenerate  
 CC oligonucleotide primers (AA046150-52) which were then used to amplify  
 CC sequences of the CB antigen (vacuolating toxin gene).  
 XX  
 XX Sequence 20 AA;  
 SQ

Query Match 0.5%; Score 7; DB 14; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 DFNAKNI 195  
 DB 3 dfnaknl 9  
 |||||

RESULT 52  
 AAY90740  
 ID AAY90740 standard; peptide; 22 AA.  
 XX  
 AC AAY90740;  
 XX  
 DT 17-AUG-2000 (first entry)  
 XX  
 DE Human BCL-2 amino acid sequence 218 to 239 SEQ ID NO:7.  
 XX  
 KW Apoptotic regulation of targeting domain; ART domain; BAX; apoptosis;  
 KW cell death; cancer; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200020446-A2.  
 XX  
 PD 13-APR-2000.  
 XX  
 XX 05-OCT-1999; 99WO-IB01680.  
 PF  
 XX 05-OCT-1998; 98US-0166028.  
 PR  
 XX

PA (UYMC-) UNIV MCGILL.  
 XX  
 PI Shore GC, Goping S;  
 XX  
 DR WPI; 2000-303740/26.  
 XX  
 PT BAX polypeptide lacking an ART domain, useful for identifying agents  
 PT that modulate apoptosis which can then be used for treating cancer -  
 XX  
 PS Example 3; Page 53; 53pp; English.  
 XX  
 CC The present invention describes a pure protein (PI) comprising a BAX  
 CC polypeptide lacking an apoptotic regulation of targeting (ART) domain.  
 CC PI has cytostatic activity and can be used in the modulation of  
 CC apoptosis. The polypeptides and methods from the present invention are  
 CC useful for identifying compounds that modulate apoptosis which can then  
 CC be used for treating cancer. The present sequence represents a human  
 CC BCL-2 peptide sequence of amino acids 218 to 239, which is used in an  
 CC example from the present invention.  
 XX  
 SQ Sequence 22 AA;  
 XX

Query Match 0.5%; Score 7; DB 21; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 SLALVGA 22  
 DB 5 slalvga 11  
 |||||

RESULT 53  
 AAR87005  
 ID AAR87005 standard; peptide; 25 AA.  
 XX  
 AC AAR87005;  
 XX  
 XX 12-JUL-1996 (first entry)  
 DT  
 XX  
 DE Class II MHC antigen alpha-chain derived peptide DR 0101 53-77.  
 XX  
 KW Major histocompatibility complex antigen; HLA;  
 KW human lymphocyte antigen; Class II; MHC; alpha-chain; CTL;  
 KW alpha-1 helix domain; immunomodulation; cytotoxic T lymphocyte;  
 KW allogeneic cell; inhibition; sensitisation; organ transplantation;  
 KW graft versus host disease; immunosuppressant drug; diagnosis  
 KW antibody production; competitive binding assay.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9534321-A1.  
 XX  
 PD 21-DEC-1995.  
 XX  
 PF 16-JUN-1995; 95WO-US07673.  
 XX  
 PR 16-JUN-1994; 94US-0260548.  
 XX  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 XX  
 PI Clayberger C, Krensky AM;  
 XX  
 DR WPI; 1996-049420/05.  
 XX  
 XX Novel peptide(s) of the alpha subunit of class II MHC antigens  
 PT useful for modulation of cytotoxic T cell activity, esp. in  
 PT therapies associated with transplantation  
 XX  
 PS Claim 6; Page 13; 23pp; English.  
 XX  
 CC The present peptide, DR 0101, comprises amino acids 53-77 of  
 CC the Class II MHC, HLA antigen alpha-chain alpha-1 helix domain.

CC It can be used to immunomodulate cytotoxic T lymphocyte (CTL)  
 CC activity against allogeneic cells, i.e. by inhibiting CTL activity  
 CC or sensitising target cells, useful in organ transplantation, e.g.  
 CC preventing graft vs. host disease. In particular it can be used  
 CC with an immunosuppressing drug, esp. 1 comprising at least 1 copy  
 CC of the B2702.75-84 peptide from an alpha-1 helix domain of a  
 CC Class I MHC antigen. It may be joined to an immunogen for the  
 CC prodn. of antibodies, and may be linked to labels, e.g. enzymes,  
 CC fluorescers and radioisotopes, for diagnostic purposes, e.g. in  
 CC a competitive binding assay to evaluate the activity CTL  
 CC immunomodulating agents.

XX Sequence 25 AA;  
 SQ

Query Match 0.5%; Score 7; DB 17; Length 25;

Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 Db |||||  
 2 feaqal 8

RESULT 54

AAW21833  
 ID AAW21833 standard; Peptide; 25 AA.

XX AAW21833;  
 AC

XX 26-OCT-1997 (first entry)  
 DT

DE Rat RT1.Dalpha class II MHC alpha chain peptide (aa51-75).  
 XX

XX Immunosuppression; major histocompatibility complex class II;  
 KW MHC; mixed lymphocyte reaction; allorecognition; cytotoxic T cell;  
 KW alloimmunity; autoimmune disease; organ transplant;  
 KW multiple sclerosis; rheumatoid arthritis; RT1.Dalpha.

XX Rattus sp.  
 OS

XX WO9710711-A1.  
 PN

XX 27-MAR-1997.  
 PD

XX 23-SEP-1996; 96WO-US15662.  
 PF

XX 21-SEP-1995; 95US-0004117.  
 PR

XX (AUTO-) AUTOIMMUNE INC.  
 PA

XX Carpenter CB, Murphy ET, Sayegh MH;  
 PI

XX WPI; 1997-202534/18.  
 DR

XX Suppression of immune responses with major histocompatibility  
 PT complex class II peptide(s) - useful in alloimmunity, e.g. organ  
 PT transplantation, and autoimmunity, e.g. in multiple sclerosis or  
 PT rheumatoid arthritis

XX Example 8; Page 25; 55pp; English.  
 PS

XX This peptide sequence comprises amino acids 51-75 of the alpha  
 CC chain of the class II MHC from rat RT1.Dalpha. The sequence is  
 CC highly conserved across alleles and species (see also AAW21830-32).  
 CC A novel method of suppressing an immune response comprises  
 CC administering a class II MHC alpha chain, or a fragment, that can  
 CC suppress at least one (preferably all) of the following: a mixed  
 CC lymphocyte reaction or other T-cell allorecognition reaction;  
 CC generation of cytotoxic T-cells recognising an alloantigen;  
 CC lymphocyte proliferation against tissue antigen; and stimulatory  
 CC cytokine production by lymphocytes. Immune responses that can be  
 CC abated or suppressed include alloimmunity (e.g. in organ

CC transplantation) and autoimmunity (e.g. in multiple sclerosis or  
 CC rheumatoid arthritis). Immune responses can be down-regulated  
 CC specifically without the adverse effects of conventional treatments.  
 XX Sequence 25 AA;  
 SQ

Query Match 0.5%; Score 7; DB 18; Length 25;

Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 Db |||||  
 4 feaqal 10

RESULT 55

AAG02215  
 ID AAG02215 standard; Protein; 57 AA.

XX AAG02215;  
 AC

XX 06-OCT-2000 (first entry)  
 DT

XX Human secreted protein, SEQ ID NO: 6296.  
 DE

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.  
 KW Homo sapiens.

XX EP1033401-A2.  
 PN

XX 06-SEP-2000.  
 PD

XX 21-FEB-2000; 2000EP-0200610.  
 PF

XX 26-FEB-1999; 99US-0122487.  
 PR

XX (GEST ) GENSET.  
 PA

XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI

XX WPI; 2000-500381/45.  
 DR

XX N-PSDB; AAC02221.  
 DR

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX Claim 13; SEQ ID 6296; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 57 AA;  
 SQ

Query Match

Best Local Similarity 0.5%; Score 7; DB 21; Length 57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 TATTTLN 954  
 |||||  
 Db 30 tatttln 36

RESULT 56  
 AAR71467  
 ID AAR71467 standard; protein; 62 AA.  
 AC AAR71467;  
 XX 09-OCT-1995 (first entry)  
 DT Mature thermophilin 1.  
 DE Bacteriocin; antibacterial.  
 KW Streptococcus thermophilus strain CNCM I-1351.  
 OS  
 XX  
 PN W09506736-A.  
 PD 09-MAR-1995.  
 XX 24-AUG-1994; 94WO-EF02805.  
 PF 03-SEP-1993; 93CH-0002628.  
 PR (NEST ) SOC PROD NESTLE SA.  
 PA Germond JE, Marciset O, Mollet B, Germond J;  
 PI WPI; 1995-115451/15.  
 DR New bacteriocin(s) from Streptococcus thermophilus - and related  
 PT nucleic acid and strains producing them, are antibacterial useful  
 PT in food processing and cosmetics.  
 XX  
 PS Claim 1; Page 33; 44pp; English.

Two protein factors are isolated from *S. thermophilus* CNCM I-1351 strain. They are bacteriocins called thermophilin 1 and thermophilin 2 (see AAR71467 and AAR71468 respectively). Such sequences, and sequences which differ by a substitution, a deletion and/or an insertion of at least one AA are claimed. The different appearance of the inhibition halos makes it possible to suspect a different antibacterial activity between thermophilin 1 and thermophilin 2. A mol. wt. of 5800 D is revealed for thermophilin 1, and 3900 D for thermophilin 2.  
 Sequence 62 AA;

Query Match 0.5%; Score 7; DB 16; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 467 GGYALAG 473  
 |||||  
 Db 11 ggyalag 17

RESULT 57  
 AAR82540  
 ID AAR82540 standard; Protein; 67 AA.  
 AC AAR82540;  
 XX 17-APR-1996 (first entry)  
 DT Hybrid IA beta chain fragment.  
 DE Major histocompatibility complex; MHC; T-cell receptor; TCR;  
 KW autoimmune disease; immunodeficiency disease; immune response;

XX

KW immunoproliferation disease; graft-host rejection; therapy; B cell;  
 KW M12.C3; pM12-IAB-Ea.  
 XX Synthetic.

XX  
 PH Location/Qualifiers  
 FT 1..27  
 FT Region /note= "leader region"  
 FT Peptide 28..31  
 FT Peptide /note= "IA beta chain beta 1 domain fragment"  
 FT Peptide 32..48  
 FT Region /note= "IE alpha chain peptide fragment"  
 FT Domain 49..63  
 FT /note= "linker region"  
 FT 64..67  
 FT /note= "IA beta chain beta 1 domain"

XX W09523814-A1.  
 XX 08-SEP-1995.  
 XX 03-MAR-1995; 95WO-US02689.  
 XX 04-MAR-1994; 94US-0207481.  
 XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.

XX Kappler JW, Marrack P;  
 XX WPI; 1995-320543/41.  
 XX N-PSDB; AAT04275.

XX Peptide-MHC complex comprising antigenic peptide, linker and MHC  
 PT segment - useful as reagents for the treatment of diseases including  
 PT auto-immune diseases, immuno-stimulatory diseases or graft-host  
 PT rejection

XX Example 2; Page 68; 94pp; English.

XX This sequence represents a fragment of a hybrid IA beta chain. This  
 CC sequence contains a fragment of the IE alpha chain (residues 56-73), as  
 CC well as a linker. The full length hybrid coding sequence (AAR04269) was  
 CC transfected into a B cell line (M12.C3) using plasmid pM12-IAB-Ea. It  
 CC was found that the encoded sequence was expressed in these cells.  
 CC Complexes such as this may be used to regulate an immune response. The  
 CC complexes are capable of being recognised by a TCR alone or in  
 CC combination with additional MHC proteins. These complexes are useful  
 CC for therapeutic purposes and experimental purposes. They can also be  
 CC used as reagents for the treatment of diseases including autoimmune  
 CC diseases, immunodeficiency diseases, immunoproliferation diseases, and  
 CC graft-host rejection.

XX Sequence 67 AA;

Query Match 0.5%; Score 7; DB 16; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQNAL 1115  
 |||||  
 Db 34 feaqnal 40

RESULT 58  
 AAB48955  
 ID AAB48955 standard; Protein; 68 AA.  
 AC AAB48955;  
 XX 27-MAR-2001 (first entry)

XX DICE-II I-Ab-restricted MHC class II epitope-containing fusion protein.  
 DE

XX Transposable element; MHC epitope; major histocompatibility complex;  
 KW intracellular bacterial pathogen; loxp site; Cre recombinase;  
 KW insertion end; in-frame fusion; detection; antigen;  
 KW disseminated insertions of class-II epitopes; DICE-II; transposon Tn5;  
 KW I-Ab-restricted MHC class II epitope.  
 XX Synthetic.  
 OS  
 XX WO200071158-A1.  
 PN  
 XX 30-NOV-2000.  
 PD  
 XX 26-MAY-2000; 2000WO-US14687.  
 XX  
 PF 26-MAY-1999; 99US-0136210.  
 XX  
 PR (UYOR-) UNIV OREGON HEALTH SCI.  
 PA  
 XX Heffron FL, Parker DC, Ellefson DD;  
 PI WPI; 2001-031967/04.  
 XX  
 DR Transposable element for detecting an antigenic epitope of a pathogen,  
 XX comprising 5' and 3' recombining sites, nucleic acid sequences encoding  
 PT a selectable marker and major histocompatibility complex (MHC) epitope,  
 PT and an insertion end.  
 XX  
 XX Disclosure; Fig 13; 63pp; English.  
 PS  
 XX The invention relates to a novel transposable element comprising DNA  
 CC encoding a selectable marker (e.g., antibiotic resistance) located  
 CC between a 5' recombining site and a 3' recombining site (e.g., loxp  
 CC sites); DNA encoding an MHC (major histocompatibility complex) epitope  
 CC either 5' of the 5' recombining site or 3' of the 3' recombining site;  
 CC and insertion ends comprising an inverted repeat sequence at the 5' and  
 CC 3' ends of the transposable element sufficient for integration of the  
 CC transposable element. The transposable elements of the invention are able  
 CC to introduce in-frame insertions throughout the chromosome of an  
 CC intracellular bacterial pathogen. This system "tags" the bacterial gene  
 CC and resulting protein, allowing the identification of proteins  
 CC secreted across the membranes of the eukaryotic cell infected by the  
 CC bacterium. In one embodiment, the transposable elements contain an  
 CC antibiotic resistance cassette, two minimal loxp recombination sites, an  
 CC MHC class I or class II epitope, and flanking insertion ends. A  
 CC transposase, such as the Cre recombinase protein, is expressed in trans  
 CC from a plasmid, or can be included in the transposable element. The Cre  
 CC recombinase loops out the intervening sequences containing the antibiotic  
 CC resistance cassette. When the transposable element inserts within a gene,  
 CC the resolved insertion places the MHC class I or class II epitope in  
 CC frame with the gene. The transposable elements of the invention are  
 CC useful for detecting an antigenic epitope of an intracellular bacterial  
 CC pathogen, such as *Salmonella* sp., *Mycobacterium tuberculosis* and *Listeria*  
 CC monocytes. Certain embodiments of the technology, termed  
 CC "disseminated insertions of class-I epitopes" (DICE-I; DICE-II for  
 CC class II epitopes) allow the rapid and accurate identification of  
 CC proteins involved in bacterial pathogenesis so that such proteins can  
 CC be used as vaccine and drug targets. Carrier vaccines may be generated  
 CC by infecting bacteria with a transposable element of the invention  
 CC which additionally comprises an antigen associated with a disease,  
 CC preferably cancer or a viral or bacterial disease, operably linked to the  
 CC MHC epitope DNA of the transposable element. The present sequence  
 CC represents an I-Ab-restricted MHC class II epitope-containing fusion  
 CC protein encoded by the resolved sequence of a DICE-II transposable  
 CC element.  
 XX  
 XX Sequence 68 AA;

Query Match 0.5%; Score 7; DB 22; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 Db 18 feaqqal 24  
 RESULT 59  
 AAG16568  
 ID AAG16568 standard; Protein; 72 AA.  
 XX  
 AC AAG16568;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 17263.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147415.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149375.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 72;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 916 LLOTLI 922  
Db 8 llttlll 14

RESULT 60  
AAG52897

ID XX AAG52897 standard; Protein; 72 AA.  
AC AAG52897;  
XX XX  
DT 18-OCT-2000 (first entry)  
XX XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67290.  
XX XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX XX  
OS Arabidopsis thaliana.  
XX XX  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126564.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 21-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147036.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151338.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155480.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157863.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 916 LQQTLLI 922  
 |||||  
 Db 8 llqtlll 14

RESULT 61  
 AAW10492  
 ID AAW10492 standard; Protein; 81 AA.  
 XX  
 AC AAW10492;  
 XX  
 DT 11-SEP-1997 (first entry)  
 XX  
 DE Alpha region of Class II MHC DRA\*0101.

XX Soluble; fusion; major histocompatibility complex; MHC; region;  
 KW heterodimer; complex; alpha1; antigen; binding groove; tolerance;  
 KW autoantigen; disease; insulin dependent; diabetes mellitus; IDDM;  
 KW antagonist; T cell; anergy; presenting cell; DRA\*0101; Class II.  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO9640944-A2.  
 PN  
 XX 19-DEC-1996.  
 PD  
 XX 07-JUN-1996; 96WO-US10102.  
 XX  
 XX 27-OCT-1995; 95US-0005964.  
 PR 07-JUN-1995; 95US-0480002.  
 PR 07-JUN-1995; 95US-0482133.  
 PR 07-JUN-1995; 95US-0483241.  
 XX  
 XX (ANER-) ANERGEN INC.  
 PA (ZYMO) ZYMOGENETICS INC.  
 PA  
 XX Deshpande S, Gross JA, Kindsvogel W, Reich EP, Sheppard PO;  
 PI  
 XX WPI; 1997-052337/05.  
 DR N-PSDB; AAT47122.  
 DR  
 XX Novel fused major histocompatibility complex:antigenic peptide  
 PT complex - useful to induce tolerance to an autoantigen-related  
 PT disease e.g. insulin-dependent diabetes mellitus  
 XX  
 PS Example 1; Page 121; 142pp; English.  
 XX  
 CC A novel soluble fused major histocompatibility complex (MHC)  
 CC heterodimer:peptide complex, comprises DNA encoding 1st and 2nd  
 CC MHC domains, e.g. the present sequence, linked by DNA encoding a  
 CC 5-25 residue linker, and a DNA encoding an antigenic peptide able  
 CC to associate with a peptide binding groove of the MHC molecule,  
 CC linked in frame to the DNA encoding the 2nd domain by a DNA  
 CC encoding a 5-25 residue linker. The complex can be used to induce  
 CC immunological tolerance in adults susceptible to, or suffering from  
 CC an autoantigen related disease, e.g. insulin dependent diabetes  
 CC mellitus (IDDM), by antagonising the binding of particular T cells  
 CC and antigen presenting cells, to induce anergy (immunological  
 CC non-responsiveness) in the targeted T cell. As the heterodimers and  
 CC obviating antigen are permanently linked into a single chain,  
 CC obviating the requirement for complex heterodimer truncation or  
 CC formation, the complex eliminates inefficient and non-specific  
 CC peptide loading.  
 XX  
 SQ Sequence 81 AA;

Query Match 0.5%; Score 7; DB 18; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1109 FEAQAL 1115  
 |||||  
 Db 54 feaqgal 60

RESULT 62  
 AAR71469  
 ID AAR71469 standard; Protein; 85 AA.  
 XX  
 AC AAR71469;  
 XX  
 DT 09-OCT-1995 (first entry)  
 XX  
 DE Premature thermophilin 1.  
 XX  
 KW Bacteriocin; antibacterial.



XX OS Streptococcus thermophilus strain CNM I-1351.  
 XX FH Key Location/Qualifiers  
 XX FT Peptide 1..23  
 XX FT /label= Leader  
 XX FT /note= "Claimed"  
 XX PN WO9506736-A.  
 XX PD 09-MAR-1995.  
 XX PF 24-AUG-1994; 94WO-EP02805.  
 XX PR 03-SEP-1993; 93CH-0002628.  
 XX PA (NEST ) SOC PROD NESTLE SA.  
 XX PI Germond JE, Marcliset O, Mollet B, Germond J;  
 XX DR WPI; 1995-115451/15.  
 XX DR N-PSDB; AAQ85766.  
 XX PT New bacteriocin(s) from Streptococcus thermophilus - and related  
 XX PT nucleic acid and strains producing them, are antibacterial useful  
 XX PT in food processing and cosmetics.  
 XX PS Claim 12; Page 35; 44pp; English.  
 XX CC Two protein factors are isolated from S. thermophilus CNM I-1351  
 XX CC strain. They are bacteriocins called thermophilin 1 and thermophilin  
 XX CC 2 (see AAR71467 and AAR71468 respectively). PCR is carried out with the  
 XX CC aid of the two non-radioactive mixtures of the degenerate sequences  
 XX CC AAQ85767 and AAQ85768 on a chromosomal DNA prepn. from the strain CNM  
 XX CC I-1351. The former corresp. to the N-terminal part and the latter  
 XX CC to the C-terminal part of thermophilin 1. Colonies are screened  
 XX CC using the probe AAQ85769 which corresp. to a sequence encoding AAs  
 XX CC 9-47 of thermophilin 1. Sequence AAQ85766 is obtd. It comprises an  
 XX CC operon encoding two proteins having the sequences AAR71469 corresp.  
 XX CC before maturation to thermophilin 1 and AAR71470 corresp. before  
 XX CC maturation to thermophilin 2. A third ORF also starts from nt 679  
 XX CC of this sequence and should corresp. to the gene for immunity.  
 XX CC AAR71469 has a leader peptide of 23 AAs which has a Gly-Gly unit  
 XX CC characteristic of a class of bacteriocins from lactic acid  
 XX CC bacteria. AAR71470 has a leader peptide of 21 AAs which also has a  
 XX CC Gly-Gly unit.  
 XX SQ Sequence 85 AA;

Query Match 0.5%; Score 7; DB 16; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 467 GGYALAG 473  
 DB 34 ggyalag 40  
 |||||

RESULT 63  
 AAB40396  
 ID AAB40396 standard; Protein; 101 AA.  
 XX  
 XX AAB40396;  
 XX  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 XX Human ORFX ORF160 polypeptide sequence SEQ ID NO:320.  
 DE  
 XX  
 XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnerable; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200058473-A2.  
 PN  
 XX  
 XX 05-OCT-2000.  
 PD  
 XX  
 XX 31-MAR-2000; 2000WO-US08621.  
 PF  
 XX  
 XX 31-MAR-1999; 99US-0127607.  
 PR  
 XX 02-APR-1999; 99US-0127636.  
 PR  
 XX 05-APR-1999; 99US-0127728.  
 PR  
 XX 30-MAR-2000; 2000US-0540763.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX  
 XX Shinkets RA, Leach M;  
 PI  
 XX  
 XX WPI; 2000-602362/57.  
 DR  
 XX N-PSDB; AAC74605.  
 DR  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 XX Claim 11; Page 599; 5507pp; English.  
 PS  
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 XX which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 XX sequences have activities such as: cytostatic; hepatotropic; vulnerary;  
 XX antiparasitic; antiparkinsonian; neurotropic; neuroprotective;  
 XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
 XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 XX antidiabetic; hypotensive; dermatological; immunosuppressive;  
 XX antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
 XX antithyroid; and antianemic. The sequences can be used for determining  
 XX the presence of or predisposition to, or preventing or treating  
 XX pathological conditions associated with an ORFX-associated disorder. The  
 XX nucleic acids can be used to express ORFX proteins in gene therapy.  
 XX vectors. The proteins and nucleic acids may be used to treat cancers,  
 XX proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 XX graft vs host disease, cardiovascular disease, diabetes mellitus,  
 XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 XX coagulation; to inhibit thrombosis; and as a contraceptive.  
 XX SQ Sequence 101 AA;

Query Match 0.5%; Score 7; DB 21; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 1.le+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 100 LYRSLLS 106  
 DB 88 lyrslls 94  
 |||||

RESULT 64  
 AAB89830

ID AAW89830 standard; Protein; 102 AA.

XX AC AAW89830;

XX DT 18-FEB-1999 (first entry)

XX DE N-terminal sequence of clone Y104.1 protein.

XX KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;

XX KW Peptic ulcer; gastric adenocarcinoma; gastric lymphoma.

XX OS Helicobacter pylori.

XX PN W09849314-A2.

XX PD 05-NOV-1998.

XX PF 27-APR-1998; 98WO-US08487.

XX PR 14-OCT-1997; 97US-0061958.

XX PR 25-APR-1997; 97US-0045107.

XX PA (GENE-) GENELABS TECHNOLOGIES INC.

XX PI Chow TP, Fry KE, Lim MY, McAttee CP;

XX DR WPI; 1999-009433/01.

XX PT New Helicobacter pylori antigens and related nucleic acid sequences  
XX PT - useful in serological diagnosis and protective vaccines, providing  
XX PT long-lasting immune response

XX PS Claim 15; Page 122; 402pp; English.

XX CC The present sequence represents a Helicobacter pylori antigenic protein  
XX CC that is characterised by immunoreactivity with H. pylori-positive  
XX CC antisera. The proteins are highly immunogenic and induce a long-lasting  
XX CC immune response that persists even after antimicrobial treatment. In  
XX CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are  
XX CC highly sensitive and specific. The specification also describes 69  
XX CC previously unrecognised immunogenic cluster families. H. pylori antigens  
XX CC are used to detect H. pylori-specific antibodies, for diagnosing  
XX CC infection or to confirm eradication of infection, and in vaccines to  
XX CC protect against H. pylori infection and related diseases (gastritis,  
XX CC peptic ulcer, gastric adenocarcinoma/lymphoma).

XX SQ Sequence 102 AA;

Query Match 0.5%; Score 7; DB 20; Length 102;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGL 23

DB 8 lalvgl 14

RESULT 65

AAG00269

ID AAG00269 standard; Protein; 118 AA.

XX AC AAG00269;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein, SEQ ID NO: 4350.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX KW gene therapy; chromosome mapping.

XX OS Homo sapiens.

XX PI

PN

XX EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST ) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX DR N-PSDB; AAC00275.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX PS Claim 13; SEQ ID 4350; 71pp + CD-ROM; English.

XX CC The present sequence is a polypeptide encoded by one of a large number  
XX CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
XX CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
XX CC different tissues. EST sequences usually correspond mainly to the 3'  
XX CC untranslated region (UTR) of the mRNA because they are often obtained  
XX CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
XX CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
XX CC those cases where longer cDNA sequences have been obtained, the full 5'  
XX CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
XX CC ends and can therefore be used to obtain full length cDNAs and genomic  
XX CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
XX CC chromosome mapping procedures. They are used to obtain upstream  
XX CC regulatory sequences and to design expression and secretion vectors.

XX SQ Sequence 118 AA;

Query Match 0.5%; Score 7; DB 21; Length 118;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 VSGLLSW 63

DB 4 vsglisw 10

RESULT 66

AAG03754

ID AAG03754 standard; Protein; 129 AA.

XX AC AAG03754;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein, SEQ ID NO: 7835.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX KW gene therapy; chromosome mapping.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST ) GENSET.

XX PI

Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI: 2000-500381/45.  
 DR N-PSDB; AAC03760.  
 XX  
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX  
 PS Claim 13; SEQ ID 7835; 71pp + CD-ROM; English.  
 XX  
 CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.  
 XX  
 SQ Sequence 129 AA;  
 Query Match 0.5%; Score 7; DB 21; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1109 FEAQGAL 1115  
 Db 79 feaqgal 85  
 RESULT 67  
 AAB54190  
 ID AAB54190 standard; Protein; 133 AA.  
 XX  
 AC AAB54190;  
 XX  
 DT 09-MAR-2001 (first entry)  
 XX  
 DE Human pancreatic cancer antigen protein sequence SEQ ID NO:642.  
 XX  
 KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;  
 KW detection; diagnosis; identification; cytostatic; neuroprotective;  
 KW nontropic; immunomodulatory; relaxant; contraceptive; gynaecological;  
 KW antiinflammatory; cardiant; gene therapy; chromosome mapping;  
 KW linkage analysis; tissue identification; tissue typing; forensic;  
 KW neural; immune system; muscular; reproductive; gastrointestinal;  
 KW pulmonary; cardiovascular; renal; proliferative.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055320-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US05989.  
 XX  
 XX 12-MAR-1999; 99US-0124270.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM;  
 XX  
 XX WPI: 2000-579444/54.  
 DR N-PSDB; AAC98955.  
 DR  
 XX New nucleic acid that is a pancreatic cancer antigen for preventing,  
 PT treating, or ameliorating a medical condition, particular pancreatic

PT cancer, or for use in assays for diagnosing a pathological condition -  
 XX Claim 11; Page 1079-1080; 1379pp; English.  
 XX  
 CC AAC98773 to AAC99231 encode the human pancreatic cancer associated  
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to  
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,  
 CC neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive,  
 CC gynaecological, cardiant and antiinflammatory activities, and can be used  
 CC in gene therapy. The polynucleotide and proteins can be used for  
 CC preventing, treating, or ameliorating a medical condition or in assays  
 CC for diagnosing a pathological condition or a susceptibility to one in a  
 CC subject. Binding partners to the proteins and the activity of the  
 CC proteins can be identified. The pancreatic cancer antigens can be used to  
 CC detect, treat or prevent pancreatic disorders, especially cancer.  
 CC Agonists and antagonists to the antigens can be screened for. The  
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic  
 CC acid hybridisation probes that can be used in chromosome mapping, linkage  
 CC analysis, tissue identification and/or typing and a variety of forensic  
 CC and diagnostic methods. The proteins can be used to generate antibodies  
 CC which are used to purify, detect and target the polypeptides, including  
 CC both in vivo and in vitro diagnostic and therapeutic methods. The  
 CC proteins can be used to treat or prevent neural, immune system, muscular,  
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or  
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 133 AA;  
 Query Match 0.5%; Score 7; DB 21; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 71 ANKTPDK 77  
 Db 36 anktpdk 42  
 RESULT 68  
 AAW55658  
 ID AAW55658 standard; Protein; 134 AA.  
 XX  
 AC AAW55658;  
 XX  
 DT 24-JUN-1998 (first entry)  
 XX  
 DE H. pylori ORF 11ae80818\_7290627\_c2\_51 cell envelope OMP.  
 XX  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9737044-A1.  
 XX  
 PD 09-OCT-1997.  
 XX  
 PF 27-MAR-1997; 97WO-US05223.  
 XX  
 PR 06-DEC-1996; 96US-0761318.  
 PR 29-MAR-1996; 96US-0625811.  
 PR 02-APR-1996; 96US-0758731.  
 PR 25-OCT-1996; 96US-0736905.  
 PR 28-OCT-1996; 96US-0738859.  
 XX  
 XX (ASTR ) ASTRA AB.  
 XX  
 XX Alm RA, Smith D;  
 XX  
 XX WPI: 1997-503122/46.  
 DR N-PSDB; AAV25067.

XX Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX  
 PS Claims 14,80; Pages 863-864; 1145pp; English.  
 XX  
 CC This is the sequence of a H. pylori cell envelope outer membrane protein  
 CC having a terminal Phe residue and a C-terminal tyrosine cluster motif.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.  
 XX  
 SQ Sequence 134 AA;

Query Match 0.5%; Score 7; DB 18; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 405 TNADGTI 411  
 |||||  
 Db 13 tnadgti 19

RESULT 69  
 AAG00183  
 ID AAG00183 standard; Protein: 145 AA.  
 AC AAG00183;  
 XX  
 DT 06-OCT-2000 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID NO: 4264.  
 XX  
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 XX EP1033401-A2.  
 PN  
 XX  
 XX 06-SEP-2000.  
 PD  
 XX  
 XX 21-FEB-2000; 2000EP-0200610.  
 PF  
 XX 26-FEB-1999; 99US-0122487.  
 PR  
 XX (GEST ) GENSET.  
 PA  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI  
 XX WPI; 2000-500381/45.  
 DR  
 XX N-PSDB; AAC00189.  
 XX  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX  
 PS Claim 13; SEQ ID 4264; 71pp + CD-ROM; English.  
 XX  
 CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.  
 XX  
 SQ Sequence 145 AA;

Query Match 0.5%; Score 7; DB 21; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 FEAQGAL 1115  
 |||||  
 Db 79 feaqqal 85

RESULT 70  
 AAW55229  
 ID AAW55229 standard; Protein: 152 AA.  
 XX  
 AC AAW55229;  
 XX  
 DT 26-JUN-1998 (first entry)  
 XX  
 DE H. pylori ORF 03xell1215orf7 protein.  
 XX

KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;  
 KW bacterium.

OS Helicobacter pylori.

XX WO9737044-A1.

XX 09-OCT-1997.

XX 27-MAR-1997; 97WO-US05223.

XX 06-DEC-1996; 96US-0761318.

XX 29-MAR-1996; 96US-0625811.

XX 02-APR-1996; 96US-0758731.

XX 25-OCT-1996; 96US-0736905.

XX 28-OCT-1996; 96US-0738859.

XX (ASTR ) ASTRA AB.

XX Alm RA, Smith D;

XX WPI; 1997-503122/46.

XX N-PSDB; AAV24636.

XX Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX

PS Claim 14; Page 475; 1145pp; English.

XX This sequence is a Helicobacter pylori protein of unspecified

CC function.  
 CC The protein may be used in a vaccine to prevent or treat *H. pylori*  
 CC infection or to identify *H. pylori* polypeptide binding compounds,  
 CC useful as potential *H. pylori* life cycle activators or inhibitors.  
 CC The DNA and probes derived from it may be used for the  
 CC identification of *H. pylori* in a sample, and the diagnosis of  
 CC *H. pylori* infection. Nucleic acid sequences complementary to the  
 CC DNA act as antisense sequences, and can be used to prevent the  
 CC translation of *H. pylori* mRNA. Antibodies against the protein can  
 CC be used in immunoassays to evaluate the abundance and distribution  
 CC of *H. pylori*-specific antigens. The genomic sequence of *H. pylori*  
 CC (ATCC 55679) was determined from overlapping contigs generated by  
 CC mechanically shearing the bacterial DNA. The sequences were  
 CC analysed for ORF of at least 180 nucleotides, and the predicted  
 CC coding regions defined by computer evaluation. To identify likely  
 CC *H. pylori* antigens for vaccine development, the amino acid  
 CC sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having  
 CC identified and determined the sequences of interest, particular  
 CC regions can be isolated from *H. pylori* by PCR amplification for  
 CC recombinant polypeptide production, e.g. in *E. coli* hosts.  
 CC  
 XX Sequence 152 AA;  
 SQ

Query Match 0.5%; Score 7; DB 18; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGL 23  
 Db 13 lalvgl 19  
 |||||

RESULT 71  
 AAY59700  
 ID AAY59700 standard; Protein; 160 AA.  
 AC AAY59700;  
 DT 18-JAN-2000 (first entry)  
 DE Secreted protein 77-8-4-P9-FL1.  
 KW Secreted protein; fingerprint identification technique;  
 KW chromosome mapping; human; hereditary disease; diagnosis; cancer;  
 KW hyperlipidaemia; cardiovascular; neurodegenerative disorder; therapy;  
 KW autoimmune disease; rheumatic disease; embryogenic disorder; myopathy;  
 KW renal injury; amino aciduria; hypoglycaemia; male rat infertility;  
 KW hypertension.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 XX Peptide 15..27  
 XX Protein 28..160  
 XX /label= signal\_peptide  
 XX /label= mature\_protein  
 XX WO9940189-A2.  
 XX 12-AUG-1999.  
 XX 09-FEB-1999; 99WO-IB00282.  
 XX 09-FEB-1998; 98US-0074121.  
 XX 13-APR-1998; 98US-0081563.  
 XX 10-AUG-1998; 98US-0096116.  
 XX 04-SEP-1998; 98US-0099273.  
 XX (GEST ) GENSET.  
 XX Bougueret L, Duclert A, Dumas Milne Edwards J;  
 XX WPI: 1999-600966/51.  
 XX N-PSDB; AA240828.  
 XX Extended cDNAs useful for expressing secreted proteins and to obtain  
 XX specific antibodies -  
 PT

XX Claim 10; Page 232; 244pp; English.  
 PS

CC This sequence represents a human secreted protein of the invention.  
 CC The extended cDNAs (or genomic DNAs obtainable from them) may be used to  
 CC prepare PCR primers and probes. These are useful for forensic matching or  
 CC positive identification by DNA sequencing. They may also be used in  
 CC alternative fingerprint identification techniques. Antibodies against the  
 CC proteins encoded by the extended cDNAs are useful in identification of  
 CC tissue types or cell species, as well as identifying tissue specific  
 CC soluble proteins. The sequences can be used for chromosome mapping and  
 CC identification of genes associated with hereditary diseases or drug  
 CC response. Signal sequences from the cDNAs can be used in construction of  
 CC secretion vectors. Other sequences derived from the extended cDNAs can be  
 CC used to clone upstream genomic DNA sequences including promoters. This is  
 CC in turn useful for identifying proteins that interact with promoter  
 CC sequences. Some of the proteins may be useful in diagnosing and treating  
 CC several disorders including, but not limited to: cancer, hyperlipidaemia,  
 CC cardiovascular and neurodegenerative disorders, autoimmune diseases, and  
 CC rheumatic diseases, embryogenic disorders, hypertension, renal injury,  
 CC amino acidurias, hypoglycaemia, male rat infertility and myopathies.  
 CC  
 XX Sequence 160 AA;  
 SQ

Query Match 0.5%; Score 7; DB 20; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 VSGLLSW 63  
 Db 4 vsllsw 10  
 |||||

RESULT 72  
 AAB23612  
 ID AAB23612 standard; Protein; 160 AA.  
 AC AAB23612;  
 DT 12-JAN-2001 (first entry)  
 DE Human secreted protein SEQ ID NO: 24.  
 KW Human; secreted protein; cytokine; cell proliferation;  
 KW nutritional supplement; immune modulation; autoimmune disorder;  
 KW haematopoiesis regulation; tissue growth; haemostasis; inflammation.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 XX Peptide 15..27  
 XX Protein 28..160  
 XX /label= signal\_peptide  
 XX /label= mature\_protein  
 XX WO200049134-A1.  
 XX 24-AUG-2000.  
 XX 18-FEB-2000; 2000WO-US04340.  
 XX 19-FEB-1999; 99US-0120680.  
 XX 23-APR-1999; 99US-0298733.  
 XX 17-AUG-1999; 99US-0149639.  
 XX 23-SEP-1999; 99US-0155686.  
 XX 01-OCT-1999; 99US-0157247.  
 XX 29-NOV-1999; 99US-0167822.  
 XX 29-NOV-1999; 99US-0167823.  
 XX 15-FEB-2000; 2000US-0298733.  
 XX (ALPH-) ALPHAGENE INC.  
 XX

PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;  
 XX WPI: 2000-549267/50.  
 DR N-PSDB; AA93112.  
 XX  
 XX New secreted proteins and polynucleotides encoding them, which are  
 PT derived from Homo sapiens, useful for therapy, diagnosis, and research,  
 PT as well as nutritional sources or supplements -  
 XX  
 XX Claim 33; Page 257-258; 309pp; English.  
 PS  
 XX The present sequence is the sequence of a human secreted protein. Its  
 CC cDNA was isolated from an adult pancreas cDNA library. The proteins  
 CC and coding sequences of the invention can be used in the isolation of  
 CC similar genes and proteins, in the elucidation of their function in vivo,  
 CC and to treat a number of conditions. It is possible that they may have  
 CC uses as nutritional supplements, as cytokine or cell proliferation  
 CC factors, in immune modulation, where they may be used to treat immune and  
 CC autoimmune diseases, as haematopoiesis regulators (treating myeloid or  
 CC lymphoid cell deficiencies), in the promotion of tissue growth, they may  
 CC have chemokine or chemotactic activity, haemostatic or thrombolytic  
 CC activity, or anti-inflammatory activity.  
 XX  
 XX Sequence 160 AA;  
 SQ  
 Query Match 0.5%; Score 7; DB 21; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 57 VSGLLSW 63  
 Db | | | | | | |  
 4 vsdllsw 10  
 RESULT 73  
 AAY87357  
 ID AAY87357 standard; Protein; 160 AA.  
 XX  
 AC AAY87357;  
 DT 11-MAY-2000 (first entry)  
 DE Human signal peptide containing protein HSPP-134 SEQ ID NO:134.  
 XX  
 XX Human; signal peptide-containing protein; HSPP; diagnosis; cancer;  
 KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;  
 KW antimicrobial; neurotropic; neuroprotective; cardiovascular; hepatotropic;  
 KW antiasthmatic; gene therapy; cell proliferation; neurological disorder;  
 KW reproductive disorder; developmental disorder; arteriosclerosis;  
 KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;  
 KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;  
 KW Parkinson's disease; Huntington's disease; ovulatory defect;  
 KW muscular dystrophy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200000610-A2.  
 PN  
 XX  
 PD 06-JAN-2000.  
 XX  
 XX 25-JUN-1999; 99WO-US14484.  
 PF  
 XX 26-JUN-1998; 98US-0090762.  
 PR 31-JUL-1998; 98US-0094983.  
 PR 01-OCT-1998; 98US-0102686.  
 PR 11-DEC-1998; 98US-0112129.  
 XX  
 XX (INCY-) INCYTE PHARM INC.  
 PA  
 XX Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;  
 PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;  
 PI Bandman O;

XX WPI: 2000-160673/14.  
 DR N-PSDB; AA298242.  
 XX  
 XX New human signal peptide-containing proteins useful in treatment,  
 PT prevention and diagnosis of e.g. cancer, inflammation and  
 PT cardiovascular disease -  
 XX  
 XX Claim 1; Page 251-252; 327pp; English.  
 PS  
 XX AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the  
 CC human signal peptide-containing proteins HSPP-1 to HSPP-134. HSPPs have  
 CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,  
 CC neuroprotective, cardiovascular and antiasthmatic activities, and can  
 CC be used in gene therapy. HSPPs can be used to treat or prevent disorders  
 CC associated with decreased activity or function of HSPP. Antagonists of  
 CC HSPP are used to treat or prevent disorders associated with increased  
 CC activity or function of HSPP. Such diseases include cell proliferation  
 CC (including cancer), inflammation, cardiovascular, neurological,  
 CC reproductive or developmental disorders, (e.g. arteriosclerosis,  
 CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,  
 CC asthma, Crohn's disease, Alzheimer's, Parkinson's or Huntington's  
 CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPP  
 CC nucleic acids can be used for the recombinant production of HSPP, for  
 CC detecting HSPP in standard hybridisation and amplification assays (for  
 CC diagnosis and monitoring), in gene therapy, as antisense,  
 CC triplex-forming or ribozyme therapeutics, for detecting related sequences  
 CC or genetic variations, and for chromosomal mapping. HSPP are also used to  
 CC raise specific antibodies (Ab) and to screen for agonists and  
 CC antagonists (potential therapeutic agents). Ab are used to diagnose, or  
 CC monitor, HSPP-related diseases (in usual immunoassays), as therapeutic  
 CC antagonists, in competitive drug screens, and for purification of HSPP  
 CC from natural sources.  
 XX  
 XX Sequence 160 AA;  
 SQ  
 Query Match 0.5%; Score 7; DB 21; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 57 VSGLLSW 63  
 Db | | | | | | |  
 4 vsdllsw 10  
 RESULT 74  
 AAY27684  
 ID AAY27684 standard; Protein; 161 AA.  
 XX  
 AC AAY27684;  
 XX  
 XX 30-JUL-1999 (first entry)  
 DT  
 XX Human secreted protein encoded by gene No. 118.  
 DE  
 XX  
 XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9924836-A1.  
 PN  
 XX 20-MAY-1999.  
 PD  
 XX 04-NOV-1998; 98WO-US23435.  
 PF

XX 17-NOV-1997; 97US-0066100.  
 PR 07-NOV-1997; 97US-0064900.  
 PR 07-NOV-1997; 97US-0064908.  
 PR 07-NOV-1997; 97US-0064911.  
 PR 07-NOV-1997; 97US-0064912.  
 PR 07-NOV-1997; 97US-0064983.  
 PR 07-NOV-1997; 97US-0064984.  
 PR 07-NOV-1997; 97US-0064985.  
 PR 07-NOV-1997; 97US-0064987.  
 PR 07-NOV-1997; 97US-0064988.  
 PR 07-NOV-1997; 97US-0066090.  
 PR 07-NOV-1997; 97US-0066094.  
 PR 07-NOV-1997; 97US-0066095.  
 PR 07-NOV-1997; 97US-0066099.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Carter KC, Ebner R, Endress CA, Feng P, Janat F;  
 PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;  
 PI Ruben SM, Shi Y, Soppet DR, Wei Y;  
 XX WPI; 1999-337740/28.  
 DR N-PSDB; AAX85050.  
 XX New human secreted proteins and coding sequences useful for treating  
 PT disorders of the immune system and hyperproliferative disorders  
 PT  
 XX Claim 11; Page 412; 507pp; English.  
 PS This sequence represents a secreted human protein encoded by the gene  
 CC clone detailed in the descriptor line. The gene can be used to generate  
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc  
 CC portion (e.g. AAX84924) for increasing the stability of the fused  
 CC protein as compared to the human protein only.  
 CC The invention relates to 125 novel genes and their fragments (nucleic  
 CC acid sequences: AAX84933-X85057; amino acid sequences AAX27367-Y27933)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 125  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAX84933 for described uses).  
 XX Sequence 161 AA;  
 SQ

Query Match 0.5%; Score 7; DB 20; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 VSGLLSW 63  
 Db 4 vsllslw 10  
 |||||

RESULT 75  
 AAB25422  
 ID AAB25422 standard; Protein; 173 AA.  
 XX  
 AC AAB25422;  
 XX  
 DT 27-NOV-2000 (first entry)  
 XX Pinus radiata cell signalling involved protein SEQ ID NO:741.  
 DE  
 DE Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
 KW environmental change; development; cell proliferation; differentiation;  
 KW elongation; survival; disease resistance; nutrient metabolism.  
 XX Pinus radiata.  
 OS

XX WO200042171-A1.  
 PN 20-JUL-2000.  
 XX 11-JAN-2000; 2000WO-US00724.  
 PF 12-JAN-1999; 99US-0228986.  
 XX 01-NOV-1999; 99US-0162866.  
 PR (GENE-) GENESIS RES & DEV CORP LTD.  
 PR Strabala TJ, Nieuwenhuizen NJ;  
 XX WPI; 2000-476052/41.  
 DR Isolated polynucleotide encoding a polypeptide involved in cell  
 XX signaling used for generating transgenic plants with modified responses  
 PT to external signals -  
 PT  
 XX Claim 3; Page 343; 527pp; English.  
 PS AAX79263 to AAX79736 and AAB25100 to AAB25570 represent polynucleotide  
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences  
 CC are involved in cell signalling. The polynucleotide and protein  
 CC sequences can be used to modify the response of plant cells to external  
 CC signals e.g. environmental changes or pathogens during the growth and  
 CC development of a plant. They can be used to modify cell proliferation,  
 CC differentiation, elongation and survival, resistance to disease and  
 CC nutrient metabolism. Examples of modifications which can be produced are  
 CC altered fruit ripening and senescence of leaves and flowers e.g. to  
 CC delay senescence and prolong the life of cut flowers or enhance  
 CC senescence of reproductive organs to engineer sterile plants. Other  
 CC modifications can be used to delay senescence in selected cell types or  
 CC organs providing fruit and vegetables which have a longer shelf life  
 CC between harvest and consumption, or to decrease branching frequency in  
 CC forest tree species giving long stretches of valuable knot-free clear  
 CC wood which can be used in solid timber furniture and veneers.  
 XX Sequence 173 AA;  
 SQ

Query Match 0.5%; Score 7; DB 21; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 966 IQTLSLS 972  
 Db 35 Iqtlsls 41  
 |||||

RESULT 76  
 AAY06802  
 ID AAY06802 standard; Protein; 174 AA.  
 XX  
 AC AAY06802;  
 XX  
 DT 23-JUN-1999 (first entry)  
 XX Peptide Seq ID No: 22.  
 DE  
 DE Major histocompatibility complex; MHC; antigen binding cleft; allergy;  
 KW antigen-specific T-cell; transplant rejection; autoimmune disease;  
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW insulin-dependent diabetes mellitus; immune response; cancer.  
 XX Homo sapiens.  
 OS  
 XX WO9914236-A1.  
 XX 25-MAR-1999.  
 PD  
 XX

PF 15-SEP-1998; 98WO-US18244.  
 XX  
 PR 10-OCT-1997; 97US-0064555.  
 PR 16-SEP-1997; 97US-0064552.  
 XX  
 PA (BURR/) BURROWS G G.  
 PA (VAND/) VANDENBARK A A.  
 XX  
 PI Burrows GG, Vandenbark AA;  
 XX WPI; 1999-229498/19.  
 DR

XX  
 PT Recombinant 2-domain major histocompatibility complex molecules used  
 PT to manipulate antigen specific T-cells  
 XX

PS Example 1; Page 68; 73pp; English.  
 XX

CC The invention relates to recombinant polypeptides comprising only those  
 CC domains of major histocompatibility complex (MHC) molecules that define  
 CC the antigen binding cleft. These polypeptides can be used to mimic the  
 CC function of mammalian MHC. The recombinant polypeptide comprises  
 CC covalently linked first and second domains, where the first domain is a  
 CC mammalian MHC class II beta domain and the second domain is a mammalian  
 CC class I alpha domain, or where the first domain is a mammalian MHC  
 CC class I alpha domain and the second domain is a mammalian MHC class I  
 CC alpha2 domain, and wherein the polypeptide is not a whole MHC class I  
 CC alpha chain. The amino terminus of the second domain is covalently linked  
 CC to the carboxy terminus of the first domain in both cases, optionally  
 CC with a peptide linker sequence. The recombinant MHC polypeptides are  
 CC useful for a wide range of in vitro and in vivo applications, and may be  
 CC used in place of either intact, purified MHC molecules or antigen  
 CC presenting cells that express MHC molecules. The polypeptides can be used  
 CC for the detection, quantification and purification of antigen-specific  
 CC T-cells in biological samples. They can also be used to activate or  
 CC inactivate T-cells and to induce anergy in T-cells. The polypeptides may  
 CC be used for the amelioration of conditions mediated by antigen-specific  
 CC T-cells, e.g. allergies, transplant rejection and autoimmune diseases  
 CC including multiple sclerosis, rheumatoid arthritis, systemic lupus  
 CC erythematosus and insulin-dependent diabetes mellitus. The polypeptides  
 CC may also be used to boost immune responses in certain conditions such as  
 CC cancer and infectious diseases. If a toxic molecule is attached to the  
 CC polypeptides, then they can also be used to kill specific T-cells.  
 XX

Sequence 174 AA;

Query Match 0.5%; Score 7; DB 20; Length 174;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;

QY 1109 FEAQAL 1115  
 Db 144 feaqal 150

RESULT 77

AAV06803  
 ID AAV06803 standard; Protein; 174 AA.  
 XX  
 AC AAV06803;

XX  
 XX  
 DT 23-JUN-1999 (first entry)  
 XX  
 DE Peptide Seq ID No: 23.

XX Major histocompatibility complex; MHC; antigen binding cleft; allergy;  
 KW antigen-specific T-cell; transplant rejection; autoimmune disease;  
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW insulin-dependent diabetes mellitus; immune response; cancer.  
 XX

OS Rattus sp.

XX W09914236-A1.  
 PN

XX  
 PD 25-MAR-1999.  
 XX  
 PF 15-SEP-1998; 98WO-US18244.  
 XX  
 PR 10-OCT-1997; 97US-0064555.  
 PR 16-SEP-1997; 97US-0064552.  
 XX  
 PA (BURR/) BURROWS G G.  
 PA (VAND/) VANDENBARK A A.  
 XX  
 PI Burrows GG, Vandenbark AA;  
 XX WPI; 1999-229498/19.  
 DR

XX  
 PT Recombinant 2-domain major histocompatibility complex molecules used  
 PT to manipulate antigen specific T-cells  
 XX

PS Example 1; Page 68-69; 73pp; English.  
 XX

CC The invention relates to recombinant polypeptides comprising only those  
 CC domains of major histocompatibility complex (MHC) molecules that define  
 CC the antigen binding cleft. These polypeptides can be used to mimic the  
 CC function of mammalian MHC. The recombinant polypeptide comprises  
 CC covalently linked first and second domains, where the first domain is a  
 CC mammalian MHC class II beta domain and the second domain is a mammalian  
 CC class I alpha domain, or where the first domain is a mammalian MHC  
 CC class I alpha domain and the second domain is a mammalian MHC class I  
 CC alpha2 domain, and wherein the polypeptide is not a whole MHC class I  
 CC alpha chain. The amino terminus of the second domain is covalently linked  
 CC to the carboxy terminus of the first domain in both cases, optionally  
 CC with a peptide linker sequence. The recombinant MHC polypeptides are  
 CC useful for a wide range of in vitro and in vivo applications, and may be  
 CC used in place of either intact, purified MHC molecules or antigen  
 CC presenting cells that express MHC molecules. The polypeptides can be used  
 CC for the detection, quantification and purification of antigen-specific  
 CC T-cells in biological samples. They can also be used to activate or  
 CC inactivate T-cells and to induce anergy in T-cells. The polypeptides may  
 CC be used for the amelioration of conditions mediated by antigen-specific  
 CC T-cells, e.g. allergies, transplant rejection and autoimmune diseases  
 CC including multiple sclerosis, rheumatoid arthritis, systemic lupus  
 CC erythematosus and insulin-dependent diabetes mellitus. The polypeptides  
 CC may also be used to boost immune responses in certain conditions such as  
 CC cancer and infectious diseases. If a toxic molecule is attached to the  
 CC polypeptides, then they can also be used to kill specific T-cells.  
 XX

Sequence 174 AA;

Query Match 0.5%; Score 7; DB 20; Length 174;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;

QY 1109 FEAQAL 1115  
 Db 144 feaqal 150

RESULT 78

AAG00963  
 ID AAG00963 standard; Protein; 176 AA.  
 XX  
 AC AAG00963;

XX  
 XX  
 DT 06-OCT-2000 (first entry)  
 XX

DE Human secreted protein, SEQ ID NO: 5044.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.  
 XX

OS Homo sapiens.



PN EPI033401-A2.  
XX 06-SEP-2000.  
PD 21-FEB-2000; 2000EP-0200610.  
PF 26-FEB-1999; 99US-0122487.  
XX (GEST ) GENSET.  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
PI WPI; 2000-500381/45.  
XX N-PSDB; AAC00969.  
DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX Claim 13; SEQ ID 5044; 71pp + CD-ROM; English.  
XX The present sequence is a polypeptide encoded by one of a large number  
XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
XX were prepared from total human RNAs or polyA+ RNAs derived from 30  
XX different tissues. EST sequences usually correspond mainly to the 3'  
XX untranslated region (UTR) of the mRNA because they are often obtained  
XX from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
XX isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
XX those cases where longer cDNA sequences have been obtained, the full 5'  
XX UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
XX ends and can therefore be used to obtain full length cDNAs and genomic  
XX DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
XX chromosome mapping procedures. They are used to obtain upstream  
XX regulatory sequences and to design expression and secretion vectors.  
XX Sequence 176 AA;  
SQ

Query Match 0.5%; Score 7; DB 21; Length 176;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 702 AMMFND 708  
DB 111 ammfnd 117

RESULT 79  
AAW10493  
ID AAW10493 standard; Protein; 207 AA.  
XX AAW10493;  
AC AAW10493;  
XX 11-SEP-1997 (first entry)  
DT Soluble fused MHC heterodimer:peptide complex pLJ13.  
DE Soluble; fusion; major histocompatibility complex; MHC;  
XX heterodimer; complex; antigen; binding groove; tolerance;  
KW autoantigen; disease; insulin dependent; diabetes mellitus; IDDM;  
KW antagonist; T cell; anergy; presenting cell.  
XX Homo sapiens.  
OS Synthetic.  
XX

Key Location/Qualifiers  
FH Region 1..21  
FT /note= "human myelin basic protein residues 82-102"  
FT Region 22..28  
FT /note= "peptide linker"  
FT Region 29..121  
FT /note= "beta1 region of Class II MHC DR1beta\*1501"  
FT Region 122..126

FT /note= "peptide linker"  
FT 127..207  
XX /note= "alpha1 region of Class II MHC DRA\*0101"  
PN W09640944-A2.  
XX 19-DEC-1996.  
PD 07-JUN-1996; 96WO-US10102.  
PF 27-OCT-1995; 95US-0005964.  
XX 07-JUN-1995; 95US-0480002.  
PR 07-JUN-1995; 95US-0482133.  
XX 07-JUN-1995; 95US-0483241.  
XX (ANER-) ANERGEN INC.  
PA (ZYMO ) ZYMOGENETICS INC.  
XX Deshpande S, Gross JA, Kindsvogel W, Reich EP, Sheppard PO;  
PI WPI; 1997-052337/05.  
XX N-PSDB; AAT47123.  
DR Novel fused major histocompatibility complex:antigenic peptide  
XX complex - useful to induce tolerance to an autoantigen related  
XX disease e.g. insulin-dependent diabetes mellitus  
XX Example 1; Pages 118-119; 142pp; English.  
XX The present sequence is a novel soluble fused major  
XX histocompatibility complex (MHC) heterodimer:peptide complex,  
XX comprising 1st and 2nd MHC domains, linked by a 5-25 residue  
XX linker, and an antigenic peptide able to associate with a peptide  
XX binding groove of the MHC molecule, linked in frame to the 2nd  
XX domain by a 5-25 residue linker. The complex can be used to induce  
XX immunological tolerance in adults susceptible to, or suffering from  
XX an autoantigen related disease, e.g. insulin dependent diabetes  
XX mellitus (IDDM), by antagonising the binding of particular T cells  
XX and antigen presenting cells, to induce anergy (immunological  
XX non-responsiveness) in the targeted T cell. As the heterodimers and  
XX corresponding antigen are permanently linked into a single chain,  
XX obviating the requirement for complex heterodimer truncation or  
XX formation, the complex eliminates inefficient and non-specific  
XX peptide loading.  
XX Sequence 207 AA;  
SQ

Query Match 0.5%; Score 7; DB 18; Length 207;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1109 FEAQAGL 1115  
DB 180 feaqagl 186

RESULT 80  
AAW20276  
ID AAW20276 standard; protein; 207 AA.  
XX AAW20276;  
AC AAW20276;  
XX 09-JUL-1997 (first entry)  
DT H. pylori inner membrane protein, 23915877.aa.  
XX Cytoplasmic; vaccine; prevention; treatment; identification;  
XX binding compound; bacterium; life cycle; activator; inhibitor;  
XX duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
OS Helicobacter pylori.  
XX



PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999;	99US-0121825.	99US-0142390.
PR	05-MAR-1999;	99US-0123180.	99US-0142803.
PR	09-MAR-1999;	99US-0123548.	99US-0142920.
PR	23-MAR-1999;	99US-0125788.	99US-0142977.
PR	25-MAR-1999;	99US-0126264.	99US-0143542.
PR	29-MAR-1999;	99US-0126785.	99US-0143624.
PR	01-APR-1999;	99US-0127462.	99US-0144005.
PR	06-APR-1999;	99US-0128234.	99US-0144085.
PR	08-APR-1999;	99US-0128714.	99US-0144086.
PR	16-APR-1999;	99US-0129845.	99US-0144325.
PR	19-APR-1999;	99US-0130077.	99US-0144331.
PR	21-APR-1999;	99US-0130449.	99US-0144332.
PR	23-APR-1999;	99US-0130510.	99US-0144333.
PR	28-APR-1999;	99US-0130891.	99US-0144334.
PR	30-APR-1999;	99US-0131149.	99US-0144335.
PR	30-APR-1999;	99US-0132048.	99US-0144352.
PR	04-MAY-1999;	99US-0132407.	99US-0144632.
PR	05-MAY-1999;	99US-0132484.	99US-0144884.
PR	06-MAY-1999;	99US-0132485.	99US-0144814.
PR	06-MAY-1999;	99US-0132486.	99US-0145088.
PR	07-MAY-1999;	99US-0132487.	99US-0145085.
PR	07-MAY-1999;	99US-0132863.	99US-0145087.
PR	11-MAY-1999;	99US-0134256.	99US-0145089.
PR	14-MAY-1999;	99US-0134218.	99US-0145192.
PR	14-MAY-1999;	99US-0134219.	99US-0145145.
PR	14-MAY-1999;	99US-0134221.	99US-0145218.
PR	14-MAY-1999;	99US-0134370.	99US-0145224.
PR	18-MAY-1999;	99US-0134768.	99US-0145276.
PR	19-MAY-1999;	99US-0134941.	99US-0145913.
PR	20-MAY-1999;	99US-0135124.	99US-0145918.
PR	21-MAY-1999;	99US-0135353.	99US-0145919.
PR	24-MAY-1999;	99US-0135629.	99US-0145951.
PR	25-MAY-1999;	99US-0136021.	99US-0146386.
PR	27-MAY-1999;	99US-0136392.	99US-0146388.
PR	28-MAY-1999;	99US-0136782.	99US-0146389.
PR	01-JUN-1999;	99US-0137222.	99US-0147038.
PR	03-JUN-1999;	99US-0137528.	99US-0147204.
PR	04-JUN-1999;	99US-0137502.	99US-0147302.
PR	07-JUN-1999;	99US-0137724.	99US-0147192.
PR	08-JUN-1999;	99US-0138094.	99US-0147260.
PR	10-JUN-1999;	99US-0138540.	99US-0147303.
PR	10-JUN-1999;	99US-0138847.	99US-0147416.
PR	14-JUN-1999;	99US-0139119.	99US-0147493.
PR	16-JUN-1999;	99US-0139452.	99US-0147935.
PR	15-JUN-1999;	99US-0139453.	99US-0148171.
PR	17-JUN-1999;	99US-0139492.	99US-0148319.
PR	18-JUN-1999;	99US-0139454.	99US-0148341.
PR	18-JUN-1999;	99US-0139455.	99US-0148565.
PR	18-JUN-1999;	99US-0139456.	99US-0148684.
PR	18-JUN-1999;	99US-0139457.	99US-0149368.
PR	18-JUN-1999;	99US-0139458.	99US-0149175.
PR	18-JUN-1999;	99US-0139459.	99US-0149426.
PR	18-JUN-1999;	99US-0139460.	99US-0149722.
PR	18-JUN-1999;	99US-0139461.	99US-0149723.
PR	18-JUN-1999;	99US-0139462.	99US-0149929.
PR	18-JUN-1999;	99US-0139463.	99US-0149929.
PR	18-JUN-1999;	99US-0139750.	99US-0149930.
PR	18-JUN-1999;	99US-0139763.	99US-0150566.
PR	21-JUN-1999;	99US-0139817.	99US-0150884.
PR	22-JUN-1999;	99US-0139899.	99US-0151065.
PR	23-JUN-1999;	99US-0140353.	99US-0151066.
PR	23-JUN-1999;	99US-0140354.	99US-0151080.
PR	24-JUN-1999;	99US-0140695.	99US-0151303.
PR	28-JUN-1999;	99US-0140823.	99US-0151438.
PR	29-JUN-1999;	99US-0140991.	99US-0151930.
PR	30-JUN-1999;	99US-0141287.	99US-0152363.
PR	01-JUL-1999;	99US-0141842.	99US-0153070.
PR	01-JUL-1999;	99US-0142154.	99US-0153758.
PR	02-JUL-1999;	99US-0142055.	99US-0154018.
PR			99US-0154039.
PR			99US-0154779.
PR			99US-0155139.

PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160788.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 210;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 665 SQFSNLT 671  
 Db 7 sqfsnlt 13

RESULT 83  
 AAB53450  
 ID AAB53450 standard; Protein; 214 AA.  
 XX AC  
 XX AAB53450;

09-MAR-2001 (first entry)

Human colon cancer antigen protein sequence SEQ ID NO:990.  
 Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 identification; cytostatic; cardioactive; neuroprotective; vulnary;  
 immunomodulatory; muscular; gynaecological; gastrointestinal;  
 nephrotropic; antineoplastic; antibacterial; gene therapy; wound;  
 neural disorder; immune system disorder; muscular disorder;  
 reproductive disorder; gastrointestinal disorder; renal disorder;  
 infectious disease; cardiovascular disorder.

Homo sapiens.

OS WO20005351-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05883.

XX 12-MAR-1999; 99US-0124270.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Rosen CA, Ruben SM;  
 XX WPI; 2000-587534/55.  
 DR N-PSDB; AAC98207.

XX Colon cancer associated gene sequences, referred to as colon cancer  
 PT antigens, useful for the treatment, prevention, and diagnosis of colon  
 PT disorders such as colon cancer

PS Claim 11; Page 1567-1568; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
 CC called human colon cancer antigens, given in AAB53234 to AAB54006. The  
 CC human colon cancer antigens can have cytostatic, cardioactive, muscular;  
 CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
 CC vulnary, nephrotropic, antineoplastic, antibacterial activities, and  
 CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
 CC proteins and antibodies to the proteins are useful for the prevention,  
 CC treatment and diagnosis of colon disorders, such as colon cancer. The  
 CC polynucleotides may be used in diagnostics and research, such as for  
 CC chromosome identification, and as hybridisation probes. The proteins  
 CC may also be used to prevent diseases such as neural disorders, immune  
 CC system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, wounds, renal disorders, infectious  
 CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
 CC AAB54007 represent sequences used in the exemplification of the present  
 CC invention.

XX Sequence 214 AA;

Query Match 0.5%; Score 7; DB 21; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 749 TGTNGIS 755  
 Db 36 tgnigs 42

RESULT 84  
 AAB51810  
 ID AAB51810 standard; Protein; 224 AA.  
 XX AC  
 XX AAB51810;

16-FEB-2001 (first entry)

Gene 28 human secreted protein homologous amino acid sequence #139.

Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
 antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
 neurotropic; neuroprotective; antibacterial; virucide; fungicide;  
 ophthalmological; vulnary; autoimmune disease; rheumatoid arthritis;  
 hyperproliferative disorders; cancer; cardiovascular disorder;  
 cardiac arrest; cerebrovascular disorder; nervous system disorder;  
 Alzheimer's disease; ocular disorder; wound healing; skin aging.

OS Dictyostelium discoideum.

XX WO200061625-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US08981.

XX 09-APR-1999; 99US-0128701.

PR 20-JAN-2000; 2000US-0177166.

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX Rosen CA, Ruben SM, Komatsoulis G;  
 XX WPI; 2000-619226/59.  
 XX  
 XX New nucleic acid molecules encoding 48 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives -  
 XX  
 XX Disclosure; Page 52-53; 500pp; English.  
 XX  
 XX Polynucleotide sequences AAC93422 - AAC93449 represent cDNA encoding  
 CC human secreted proteins AAB51724 - AAB51777. Sequences AAB51778 -  
 CC AAB51825 represent alternative polypeptides encoded by the genes, and  
 CC amino acid sequences to which they are homologous. The genes and proteins  
 CC have activities dependent on the tissues and cells in which they are  
 CC expressed. Examples of their activities include immunosuppressive;  
 CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
 CC vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;  
 CC virucide; fungicide; ophthalmological; and vulnerary. The secreted  
 CC proteins, polynucleotides, antagonists and agonists may be useful in  
 CC treating, preventing and/or diagnosing diseases and disorders such as  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to  
 CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
 CC components. Oligonucleotide AAC93413 - AAC93421 and peptide AAB51723 are  
 CC used in the isolation and characterisation of the proteins and  
 CC polynucleotides of the invention.  
 XX  
 XX Sequence 224 AA;  
 SQ

Query Match 0.5%; Score 7; DB 21; Length 224;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 EEQFKER 766  
 Db |||||  
 54 eeqfker 60

RESULT 85  
 AAB69576  
 ID AAB69576 standard; protein; 228 AA.  
 XX  
 AC AAB69576;  
 XX  
 DT 11-MAY-2001 (first entry)  
 XX  
 DE Swine leukocyte antigen DRA allele consensus sequence.  
 XX  
 KW Swine leukocyte antigen; SLA; SLA-DRA; immunosuppressive; vaccine;  
 KW allospecific; xenospecific; antigen specific; allopeptide;  
 KW human suppressor CD8+CD28- T cell; autoimmune disease;  
 KW immunosuppressant therapy; allograft rejection; xenograft rejection.  
 XX  
 OS Sus sp.  
 XX  
 XX WO200076320-A1.  
 XX  
 PD 21-DEC-2000.

XX 15-JUN-2000; 2000WO-US16594.  
 XX  
 XX 15-JUN-1999; 99US-0333809.  
 XX  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 XX Suciu-Foca N, Cortensini R, Liu Z, Chang C;  
 XX WPI; 2001-025327/03.  
 XX  
 XX Generating antigen specific allospecific, xenospecific and allopeptide  
 PT antigen specific human suppressor CD8+CD28- T cells, useful for  
 PT preventing allograft or xenograft rejection and preventing autoimmune  
 PT disease -  
 XX  
 XX Disclosure; Fig 29; 131pp; English.  
 XX  
 XX The present sequence is provided in a specification relating to a  
 CC method for generating antigen specific allospecific, xenospecific  
 CC and allopeptide antigen specific human suppressor CD8+CD28- T cells.  
 CC The T cells are useful for determining whether the level of  
 CC immunosuppressant therapy is too high, preventing, or reducing the risk  
 CC or level of allograft rejection in a subject undergoing immunosuppression  
 CC therapy, preventing xenograft rejection, and preventing or suppressing an  
 CC autoimmune disease. The new method produces T suppressor cells that are  
 CC useful for specific immunosuppressive therapy so there is decreased  
 CC risk of infections and malignancies compared to previous  
 CC immunosuppressive treatments. It also allows easy reproduction of T  
 CC suppressor cells in culture. The present sequence may be used for  
 CC generating xenospecific human suppressor T cells in the method of the  
 CC invention.  
 XX  
 XX Sequence 228 AA;  
 SQ

Query Match 0.5%; Score 7; DB 22; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQCAL 1115  
 Db |||||  
 53 feaqgal 59

RESULT 86  
 AAB69577  
 ID AAB69577 standard; protein; 228 AA.  
 XX  
 AC AAB69577;  
 XX  
 DT 11-MAY-2001 (first entry)  
 XX  
 DE Swine leukocyte antigen DRA allele SLADRA-0102.  
 XX  
 KW Swine leukocyte antigen; SLA; SLA-DRA; immunosuppressive; vaccine;  
 KW allospecific; xenospecific; antigen specific; allopeptide;  
 KW human suppressor CD8+CD28- T cell; autoimmune disease;  
 KW immunosuppressant therapy; allograft rejection; xenograft rejection.  
 XX  
 OS Sus sp.  
 XX  
 XX WO200076320-A1.  
 XX  
 XX 21-DEC-2000.  
 XX  
 XX 15-JUN-2000; 2000WO-US16594.  
 XX  
 XX 15-JUN-1999; 99US-0333809.  
 XX  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 XX Suciu-Foca N, Cortensini R, Liu Z, Chang C;

XX WPI: 2001-025327/03.  
 XX Generating antigen specific, allospecific, xenospecific and allopeptide  
 PT antigen specific human suppressor CD8+CD28- T cells, useful for  
 PT preventing allograft or xenograft rejection and preventing autoimmune  
 PT disease -  
 XX Disclosure; Fig 29; 131pp; English.  
 PS  
 XX The present sequence is provided in a specification relating to a  
 CC method for generating antigen specific allospecific, xenospecific  
 CC and allopeptide antigen specific human suppressor CD8+CD28- T cells.  
 CC The T cells are useful for determining whether the level of  
 CC immunosuppressant therapy is too high, preventing, or reducing the risk  
 CC or level of allograft rejection in a subject undergoing immunosuppression  
 CC therapy, preventing xenograft rejection, and preventing or suppressing an  
 CC autoimmune disease. The new method produces T suppressor cells that are  
 CC useful for specific immunosuppressive therapy so there is decreased  
 CC risk of infections and malignancies compared to previous  
 CC immunosuppressive treatments. It also allows easy reproduction of T  
 CC suppressor cells in culture. The present sequence may be used for  
 CC generating xenospecific human suppressor T cells in the method of the  
 CC invention.  
 XX Sequence 228 AA;  
 SQ

Query Match 0.5%; Score 7; DB 22; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAGAL 1115  
 DB 53 feaqal 59  
 |||||

RESULT 87  
 AAB69578  
 ID AAB69578 standard; protein: 228 AA.  
 AC  
 AC AAB69578;  
 XX  
 DT 11-MAY-2001 (first entry)  
 XX  
 XX Swine leukocyte antigen DRA allele Siadra-0202.  
 DE  
 XX Swine leukocyte antigen; SLA; SLA-DRA; immunosuppressive; vaccine;  
 KW allospecific; xenospecific; antigen specific; allopeptide;  
 KW human suppressor CD8+CD28- T cell; autoimmune disease;  
 KW immunosuppressant therapy; allograft rejection; xenograft rejection.  
 XX  
 OS Sus sp.  
 XX  
 PN WO200076320-A1.  
 XX  
 PD 21-DEC-2000.  
 XX  
 PF 15-JUN-2000; 2000WO-US16594.  
 XX  
 PR 15-JUN-1999; 99US-0333809.  
 XX  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 PA  
 XX Suciu-Foca N, Cortensini R, Liu Z, Chang C;  
 PI WPI: 2001-025327/03.  
 XX  
 XX Generating antigen specific allospecific, xenospecific and allopeptide  
 PT antigen specific human suppressor CD8+CD28- T cells, useful for  
 PT preventing allograft or xenograft rejection and preventing autoimmune  
 PT disease -  
 XX

PS Disclosure; Fig 29; 131pp; English.  
 XX  
 CC The present sequence is provided in a specification relating to a  
 CC method for generating antigen specific allospecific, xenospecific  
 CC and allopeptide antigen specific human suppressor CD8+CD28- T cells.  
 CC The T cells are useful for determining whether the level of  
 CC immunosuppressant therapy is too high, preventing, or reducing the risk  
 CC or level of allograft rejection in a subject undergoing immunosuppression  
 CC therapy, preventing xenograft rejection, and preventing or suppressing an  
 CC autoimmune disease. The new method produces T suppressor cells that are  
 CC useful for specific immunosuppressive therapy so there is decreased  
 CC risk of infections and malignancies compared to previous  
 CC immunosuppressive treatments. It also allows easy reproduction of T  
 CC suppressor cells in culture. The present sequence may be used for  
 CC generating xenospecific human suppressor T cells in the method of the  
 CC invention.  
 XX Sequence 228 AA;  
 SQ

Query Match 0.5%; Score 7; DB 22; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAGAL 1115  
 DB 53 feaqal 59  
 |||||

RESULT 88  
 AAB69579  
 ID AAB69579 standard; protein: 228 AA.  
 AC  
 AC AAB69579;  
 XX  
 DT 11-MAY-2001 (first entry)  
 XX  
 XX Swine leukocyte antigen DRA allele Siadra-0203.  
 DE  
 XX Swine leukocyte antigen; SLA; SLA-DRA; immunosuppressive; vaccine;  
 KW allospecific; xenospecific; antigen specific; allopeptide;  
 KW human suppressor CD8+CD28- T cell; autoimmune disease;  
 KW immunosuppressant therapy; allograft rejection; xenograft rejection.  
 XX  
 OS Sus sp.  
 XX  
 PN WO200076320-A1.  
 XX  
 PD 21-DEC-2000.  
 XX  
 PF 15-JUN-2000; 2000WO-US16594.  
 XX  
 PR 15-JUN-1999; 99US-0333809.  
 XX  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 PA  
 XX Suciu-Foca N, Cortensini R, Liu Z, Chang C;  
 PI WPI: 2001-025327/03.  
 XX  
 XX Generating antigen specific allospecific, xenospecific and allopeptide  
 PT antigen specific human suppressor CD8+CD28- T cells, useful for  
 PT preventing allograft or xenograft rejection and preventing autoimmune  
 PT disease -  
 XX  
 PS Disclosure; Fig 29; 131pp; English.  
 XX  
 CC The present sequence is provided in a specification relating to a  
 CC method for generating antigen specific allospecific, xenospecific  
 CC and allopeptide antigen specific human suppressor CD8+CD28- T cells.  
 CC The T cells are useful for determining whether the level of  
 CC immunosuppressant therapy is too high, preventing, or reducing the risk  
 CC or level of allograft rejection in a subject undergoing immunosuppression

CC therapy, preventing xenograft rejection, and preventing or suppressing an  
 CC autoimmune disease. The new method produces T suppressor cells that are  
 CC useful for specific immunosuppressive therapy so there is decreased  
 CC risk of infections and malignancies compared to previous  
 CC immunosuppressive treatments. It also allows easy reproduction of T  
 CC suppressor cells in culture. The present sequence may be used for  
 CC generating xenospecific human suppressor T cells in the method of the  
 CC invention.

XX SQ Sequence 228 AA;

Query Match 0.5%; Score 7; DB 22; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 |||||  
 Db 53 feaqal 59

RESULT 89.

AAB69580  
 ID AAB69580 standard; protein; 228 AA.

XX AAB69580;

AC AAB69580;

XX 11-MAY-2001 (first entry)

DT Swine leukocyte antigen DRA allele Siadra-0101.

DE Swine leukocyte antigen; SLA; SLA-DRA; immunosuppressive; vaccine;  
 XX allospecific; xenospecific; antigen specific; allopeptide;  
 KW human suppressor CD8+CD28- T cell; autoimmune disease;  
 KW immunosuppressant therapy; allograft rejection; xenograft rejection.

XX Sus sp.

OS WO200076320-A1.

XX 21-DEC-2000.

XX 15-JUN-2000; 2000WO-US16594.

XX 15-JUN-1999; 99US-0333809.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Suciu-Foca N, Cortensini R, Liu Z, Chang C;

XX WPI; 2001-025327/03.

XX Generating antigen specific allospecific, xenospecific and allopeptide  
 XX antigen specific human suppressor CD8+CD28- T cells, useful for  
 XX preventing allograft or xenograft rejection and preventing autoimmune  
 XX disease -

XX Disclosure; Fig 29; 131pp; English.

XX The present sequence is provided in a specification relating to a  
 XX method for generating antigen specific allospecific, xenospecific  
 XX and allopeptide antigen specific human suppressor CD8+CD28- T cells.  
 XX The T cells are useful for determining whether the level of  
 XX immunosuppressant therapy is too high, preventing, or reducing the risk  
 XX or level of allograft rejection in a subject undergoing immunosuppression  
 XX therapy, preventing xenograft rejection, and preventing or suppressing an  
 XX autoimmune disease. The new method produces T suppressor cells that are  
 XX useful for specific immunosuppressive therapy so there is decreased  
 XX risk of infections and malignancies compared to previous  
 XX immunosuppressive treatments. It also allows easy reproduction of T  
 XX suppressor cells in culture. The present sequence may be used for  
 XX generating xenospecific human suppressor T cells in the method of the  
 XX invention.

XX SQ Sequence 228 AA;

Query Match 0.5%; Score 7; DB 22; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 |||||  
 Db 53 feaqal 59

RESULT 90

AAB69581  
 ID AAB69581 standard; protein; 228 AA.

XX AAB69581;

AC AAB69581;

XX 11-MAY-2001 (first entry)

DT Swine leukocyte antigen DRA allele Siadra-02011.

DE Swine leukocyte antigen; SLA; SLA-DRA; immunosuppressive; vaccine;  
 XX allospecific; xenospecific; antigen specific; allopeptide;  
 KW human suppressor CD8+CD28- T cell; autoimmune disease;  
 KW immunosuppressant therapy; allograft rejection; xenograft rejection.

XX Sus sp.

OS WO200076320-A1.

XX 21-DEC-2000.

XX 15-JUN-2000; 2000WO-US16594.

XX 15-JUN-1999; 99US-0333809.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Suciu-Foca N, Cortensini R, Liu Z, Chang C;

XX WPI; 2001-025327/03.

XX Generating antigen specific allospecific, xenospecific and allopeptide  
 XX antigen specific human suppressor CD8+CD28- T cells, useful for  
 XX preventing allograft or xenograft rejection and preventing autoimmune  
 XX disease -

XX Disclosure; Fig 29; 131pp; English.

XX The present sequence is provided in a specification relating to a  
 XX method for generating antigen specific allospecific, xenospecific  
 XX and allopeptide antigen specific human suppressor CD8+CD28- T cells.  
 XX The T cells are useful for determining whether the level of  
 XX immunosuppressant therapy is too high, preventing, or reducing the risk  
 XX or level of allograft rejection in a subject undergoing immunosuppression  
 XX therapy, preventing xenograft rejection, and preventing or suppressing an  
 XX autoimmune disease. The new method produces T suppressor cells that are  
 XX useful for specific immunosuppressive therapy so there is decreased  
 XX risk of infections and malignancies compared to previous  
 XX immunosuppressive treatments. It also allows easy reproduction of T  
 XX suppressor cells in culture. The present sequence may be used for  
 XX generating xenospecific human suppressor T cells in the method of the  
 XX invention.

XX SQ Sequence 228 AA;

Query Match 0.5%; Score 7; DB 22; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1109 FEAQAL 1115  
 DB 53 feaqal 59

RESULT 91  
 AAB69582  
 ID AAB69582 standard; protein; 228 AA.  
 AC AAB69582;  
 XX  
 DT 11-MAY-2001 (first entry)  
 XX  
 DE Swine leukocyte antigen DRA allele Siadra-02012.  
 XX  
 KW Swine leukocyte antigen; SLA; SLA-DRA; immunosuppressive; vaccine;  
 KW allospecific; xenospecific; antigen specific; allopeptide;  
 KW human suppressor CD8+CD28- T cell; autoimmune disease;  
 KW immunosuppressant therapy; allograft rejection; xenograft rejection.  
 XX  
 OS Sus sp.  
 XX  
 PN WO200076320-A1.  
 XX  
 PD 21-DEC-2000.  
 XX  
 PF 15-JUN-2000; 2000WO-US16594.  
 XX  
 PR 15-JUN-1999; 99US-0333809.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Suciu-Foca N, Cortensini R, Liu Z, Chang C;  
 XX  
 DR WPI; 2001-025327/03.  
 XX  
 PT Generating antigen specific allospecific, xenospecific and allopeptide  
 PT antigen specific human suppressor CD8+CD28- T cells, useful for  
 PT preventing allograft or xenograft rejection and preventing autoimmune  
 PT disease -  
 XX  
 PS Disclosure; Fig 29; 131pp; English.  
 XX  
 CC The present sequence is provided in a specification relating to a  
 CC method for generating antigen specific allospecific, xenospecific  
 CC and allopeptide antigen specific human suppressor CD8+CD28- T cells.  
 CC The T cells are useful for determining whether the level of  
 CC immunosuppressant therapy is too high, preventing, or reducing the risk  
 CC or level of allograft rejection in a subject undergoing immunosuppression  
 CC therapy, preventing xenograft rejection, and preventing or suppressing an  
 CC autoimmune disease. The new method produces T suppressor cells that are  
 CC useful for specific immunosuppressive therapy so there is decreased  
 CC risk of infections and malignancies compared to previous  
 CC immunosuppressive treatments. It also allows easy reproduction of T  
 CC suppressor cells in culture. The present sequence may be used for  
 CC generating xenospecific human suppressor T cells in the method of the  
 CC invention.  
 XX  
 SQ Sequence 228 AA;

Query Match 0.5%; Score 7; DB 22; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1109 FEAQAL 1115  
 DB 53 feaqal 59

RESULT 92  
 AAR74037  
 ID AAR74037 standard; protein; 229 AA.

XX AAR74037;  
 AC  
 XX 29-NOV-1995 (first entry)  
 DT  
 XX MHC polypeptide HLA DR2-Dw2 alpha chain.  
 DE  
 XX Major histocompatibility complex; T cell; modulation; immunological;  
 KW disorders; autoimmune disease; allergy; transplant rejection;  
 KW rheumatoid arthritis.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9511702-A.  
 XX  
 PD 04-MAY-1995.  
 XX  
 PF 25-OCT-1994; 94WO-US12231.  
 XX  
 PR 25-OCT-1993; 93US-0143575.  
 XX  
 PA (ANER-) ANERGEN INC.  
 XX  
 PI Nag B, Rhodes ET;  
 XX  
 DR WPI; 1995-178654/23.  
 DR N-PSDB; AAQ92015.  
 XX  
 PT New recombinant, non-glycosylated MHC polypeptide - produced in  
 PT prokaryotic cells, useful for modulating T cell function esp. in  
 PT auto-immune disease  
 XX  
 PS Disclosure; Page 50; 68pp; English.  
 XX  
 CC The sequence is that of a synthetic non-glycosylated polypeptide contg.  
 CC MHC HLA DR2-Dw2 alpha chain. The peptide may be used to modulate T  
 CC cell function for the treatment of autoimmune disease, allergic  
 CC responses, transplant rejection and other immunological disorders,  
 CC especially rheumatoid arthritis.  
 CC See also AAR74038.  
 XX  
 SQ Sequence 229 AA;

Query Match 0.5%; Score 7; DB 16; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1109 FEAQAL 1115  
 DB 54 feaqal 60

RESULT 93  
 AAW01021  
 ID AAW01021 standard; protein; 229 AA.  
 XX  
 AC AAW01021;  
 XX  
 DT 18-DEC-1996 (first entry)  
 DE  
 XX Apoptosis-blocking protein Bcl-2 mutant 106-5 (del106-115).  
 KW Apoptosis-regulating protein; Bcl-2; oncogene;  
 KW adenovirus E1B 19K protein; cell death; cancer; tumour;  
 KW immune disorder; diagnosis; therapy; BiplA; Bip13; Bip5; Nip1;  
 KW Nip2; Nip3.  
 XX  
 OS Synthetic.  
 XX  
 PN EP733706-A2.  
 XX  
 PD 25-SEP-1996.



XX 21-MAR-1996; 96EP-0104542.  
 XX PF  
 XX PR  
 XX XX  
 XX 21-MAR-1995; 95US-0408095.  
 XX XX  
 XX (UYSL-) UNIV ST LOUIS.  
 XX PI  
 XX Chinnadurai G;  
 XX XX  
 XX WPI; 1996-427055/43.  
 XX DR  
 XX Nucleic acids encoding apoptosis regulating proteins - useful for  
 PT diagnosing and treating immune disorders, malignancies, etc.  
 PT PT  
 XX XX  
 XX PS  
 XX Example 8; Page 36; 60pp; English.  
 CC The 106-5 mutant (AAW01021) of the bcl-2 oncogene product (AAW01018)  
 CC lacks amino acids 106-115 of the native protein. This and other  
 CC Bcl-2 mutants (see also AAW01019-20) were used in a two hybrid assay  
 CC to examine the interactions between Bcl-2 and novel apoptosis-  
 CC regulating proteins Nip1, Nip2 and Nip3 (AAW00997-99). The Nip  
 CC proteins were unable to interact with mutant 106-5. The site of  
 CC deletion in this mutant corresponds to a motif (see also AAW01004)  
 CC on Bcl-2 essential for interaction with Nip proteins. A second  
 CC binding motif (AAW01003) of Bcl-2 was also identified, and both  
 CC show homology to motifs (AAW01005-06) found on the 19K protein  
 CC (AAW01010) of adenovirus E1B.  
 XX XX  
 XX SQ Sequence 229 AA;

Query Match 0.5%; Score 7; DB 17; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 Db 212 slalvga 218  
 |||||

RESULT 94  
 AAW94348  
 ID AAW94348 standard; Protein; 229 AA.  
 XX AC  
 XX AAW94348;  
 XX DT 13-APR-1999 (first entry)  
 XX DE Human Bcl-2 mutant protein #106-5.  
 XX KW Human; Nip1; Nip2; Nip3; Bip1a; Bip5; Bip13; adenovirus; cell death;  
 KW viral infection; Bcl-2; protooncogene; mutational analysis; apoptosis;  
 KW E1B 19K protein; cell survival regulation.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN US5858678-A.  
 XX PD 12-JAN-1999.  
 XX PF 21-MAR-1995; 95US-0408095.  
 XX PR 02-AUG-1994; 94US-0284139.  
 XX PR 21-MAR-1995; 95US-0408095.  
 XX PA (UYSL-) UNIV ST LOUIS.  
 XX PI Chinnadurai G;  
 XX WPI; 1999-152099/13.  
 XX Polypeptides that bind to anti-apoptotic proteins - useful for

PT protecting against cell death induced by viral infection and to  
 XX modulate response to physical and chemical stimuli  
 XX PS  
 XX Example 8; Column 45-48; 41pp; English.  
 CC The present invention describes: (1) a method for regulating cell death,  
 CC comprising exposing an isolated cell to a polypeptide selected from  
 CC Nip1, Nip2, Nip3, Bip1a, Bip5 and Bip13; (2) a method for neutralising  
 CC the activity of the adenovirus E1B 19 kD protein, the Bcl-2 protein or  
 CC the BHRF-1 protein, comprising exposing an isolated cell to a  
 CC polypeptide as in (1); and (3) a method for detecting molecules that  
 CC bind to at least one polypeptide as in (1), comprising lysing cells,  
 CC exposing the lysate to the polypeptide and detecting any molecule-  
 CC polypeptide aggregates. The methods are useful for providing proteins  
 CC able to bind to other proteins known to regulate cell survival e.g. it  
 CC is known that E1B 19K protein provides a survival function similar to  
 CC the cellular protooncogene bcl-2 gene product which is able to block  
 CC apoptosis in haematopoietic B and T cells. The present sequence  
 CC represents a human Bcl-2 mutant protein from the present invention.  
 XX SQ Sequence 229 AA;

Query Match 0.5%; Score 7; DB 20; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 Db 212 slalvga 218  
 |||||

RESULT 95  
 AAW01019  
 ID AAW01019 standard; Protein; 232 AA.  
 XX AC  
 XX AAW01019;  
 XX DT 18-DEC-1996 (first entry)  
 XX DE Apoptosis-blocking protein Bcl-2 mutant 42-8 (del42-48).  
 XX KW Apoptosis-regulating protein; Bcl-2; oncogene;  
 KW adenovirus E1B 19K protein; cell death; cancer;  
 KW immune disorder; diagnosis; therapy; Bip1a; Bip13; Bip5; Nip1;  
 KW Nip2; Nip3.  
 XX OS Synthetic.  
 XX PN EP733706-A2.  
 XX PD 25-SEP-1996.  
 XX PF 21-MAR-1996; 96EP-0104542.  
 XX PR 21-MAR-1995; 95US-0408095.  
 XX PA (UYSL-) UNIV ST LOUIS.  
 XX PI Chinnadurai G;  
 XX WPI; 1996-427055/43.  
 XX Nucleic acids encoding apoptosis regulating proteins - useful for  
 PT diagnosing and treating immune disorders, malignancies, etc.  
 PT PT  
 XX XX  
 XX PS  
 XX Example 8; Page 33-34; 60pp; English.  
 CC The 42-8 mutant (AAW01019) of the bcl-2 oncogene product (AAW01018)  
 CC lacks amino acids 42-48 of the native protein. This and other  
 CC Bcl-2 mutants (see also AAW01020-21) were used in a two hybrid assay  
 CC to examine the interactions between Bcl-2 and novel apoptosis-  
 CC regulating proteins Nip1, Nip2 and Nip3 (AAW00997-99). The Nip

CC proteins were unable to interact with mutant 42-8. The site of  
 CC deletion in this mutant corresponds to a motif (see also AAW01003)  
 CC on Bcl-2 essential for interaction with Nip proteins. A second  
 CC binding motif (AAW01004) of Bcl-2 was also identified, and both  
 CC show homology to motifs (AAW01005-06) found on the 19K protein  
 CC (AAW01010) of adenovirus E1B.

XX Sequence 232 AA;

Query Match 0.5%; Score 7; DB 17; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22

Db 215 slalvga 221

RESULT 96

AAW01020  
 ID AAW01020 standard; Protein; 232 AA.

AC AAW01020;

DT 18-DEC-1996 (first entry)

DE Apoptosis-blocking protein Bcl-2 mutant 80-6 (del80-86).

XX Apoptosis-regulating protein; Bcl-2; oncogene;

KW adenovirus E1B 19K protein; cell death; cancer; tumour;

KW immune disorder; diagnosis; therapy; Bipla; Bip3; Bip5; Nip1;

KW Nip2; Nip3.

XX Synthetic.

XX EP733706-A2.

XX 25-SEP-1996.

XX 21-MAR-1996; 96EP-0104542.

XX 21-MAR-1995; 95US-0408095.

XX (UYSL-) UNIV ST LOUIS.

XX Chinnadurai G;

XX WPI; 1996-427055/43.

XX Nucleic acids encoding apoptosis regulating proteins - useful for  
 PT diagnosing and treating immune disorders, malignancies, etc.

XX Example 8; Page 34-35; 60pp; English.

XX The 80-6 mutant (AAW01020) of the bcl-2 oncogene product (AAW01018)  
 CC lacks amino acids 80-86 of the native protein. This and other  
 CC Bcl-2 mutants (see also AAW01019-21) were used in a two hybrid assay  
 CC to examine the interactions between Bcl-2 and novel apoptosis-  
 CC regulating proteins Nip1, Nip2 and Nip3 (AAW00997-99). 2 Motifs  
 CC (AAW01003-04) on Bcl-2 were identified that are essential for  
 CC interaction with the Nip proteins. These motifs show homology  
 CC to motifs (AAW01005-06) identified on the adenovirus E1B 19K  
 CC apoptosis-blocking protein (AAW01010).

XX Sequence 232 AA;

Query Match 0.5%; Score 7; DB 17; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22

Db 215 slalvga 221

RESULT 97

AAW94346  
 ID AAW94346 standard; Protein; 232 AA.

XX AAW94346;

XX 13-APR-1999 (first entry)

XX Human Bcl-2 mutant protein #42-8.

XX Human; Nip1; Nip2; Nip3; Bip3; Bip5; Bip3; adenovirus; cell death;

KW viral infection; Bcl-2; protooncogene; mutational analysis; apoptosis;

KW E1B 19K protein; cell survival regulation.

XX Homo sapiens.

XX Synthetic.

XX US5858678-A.

XX 12-JAN-1999.

XX 21-MAR-1995; 95US-0408095.

XX 02-AUG-1994; 94US-0284139.

XX 21-MAR-1995; 95US-0408095.

XX (UYSL-) UNIV ST LOUIS.

XX Chinnadurai G;

XX WPI; 1999-152099/13.

XX Polypeptides that bind to anti-apoptotic proteins - useful for  
 PT protecting against cell death induced by viral infection and to  
 PT modulate response to physical and chemical stimuli

XX Example 8; Column 43-44; 41pp; English.

XX The present invention describes: (1) a method for regulating cell death,  
 CC comprising exposing an isolated cell to a polypeptide selected from  
 CC Nip1, Nip2, Nip3, Bip1A, Bip5 and Bip3; (2) a method for neutralising  
 CC the activity of the adenovirus E1B 19 kD protein, the Bcl-2 protein or  
 CC the BHRF-1 protein, comprising exposing an isolated cell to a  
 CC polypeptide as in (1); and (3) a method for detecting molecules that  
 CC bind to at least one polypeptide as in (1), comprising lysing cells,  
 CC exposing the lysate to the polypeptide and detecting any molecule-  
 CC polypeptide aggregates. The methods are useful for providing proteins  
 CC able to bind to other proteins known to regulate cell survival e.g. it  
 CC is known that E1B 19K protein provides a survival function similar to  
 CC the cellular protooncogene bcl-2 gene product which is able to block  
 CC apoptosis in haematopoietic B and T cells. The present sequence  
 CC represents a human Bcl-2 mutant protein from the present invention.

XX Sequence 232 AA;

Query Match 0.5%; Score 7; DB 20; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22

Db 215 slalvga 221

RESULT 98

AAW94347  
 ID AAW94347 standard; Protein; 232 AA.

XX

AC AAW94347;  
 XX 13-APR-1999 (first entry)  
 DT Human Bcl-2 mutant protein #80-6.  
 DE Human; Nip1; Nip2; Nip3; Bip1A; Bip5; Bip13; adenovirus; cell death;  
 KW Viral infection; Bcl-2; protooncogene; mutational analysis; apoptosis;  
 KW E1B 19K protein; cell survival regulation.  
 XX Homo sapiens.  
 OS Synthetic.  
 PN US5858678-A.  
 XX 12-JAN-1999.  
 PD 21-MAR-1995; 95US-0408095.  
 XX 02-AUG-1994; 94US-0284139.  
 PR 21-MAR-1995; 95US-0408095.  
 XX (UYSL-) UNIV ST LOUIS.  
 PA Chinnadurai G;  
 XX WPI; 1999-152099/13.  
 DR Polypeptides that bind to anti-apoptotic proteins - useful for  
 PT protecting against cell death induced by viral infection and to  
 PT modulate response to physical and chemical stimuli  
 XX Example 8; Column 43-46; 41pp; English.

XX The present invention describes: (1) a method for regulating cell death,  
 CC comprising exposing an isolated cell to a polypeptide selected from  
 CC Nip1, Nip2, Nip3, Bip1A, Bip5 and Bip13; (2) a method for neutralising  
 CC the activity of the adenovirus E1B 19 KB protein, the Bcl-2 protein or  
 CC the BHRF-1 protein, comprising exposing an isolated cell to a  
 CC polypeptide as in (1); and (3) a method for detecting molecules that  
 CC bind to at least one polypeptide as in (1), comprising lysing cells,  
 CC exposing the lysate to the polypeptide and detecting any molecule-  
 CC polypeptide aggregates. The methods are useful for providing proteins  
 CC able to bind to other proteins known to regulate cell survival e.g. it  
 CC is known that E1B 19K protein provides a survival function similar to  
 CC the cellular protooncogene bcl-2 gene product which is able to block  
 CC apoptosis in haematopoietic B and T cells. The present sequence  
 CC represents a human Bcl-2 mutant protein from the present invention.

XX Sequence 232 AA;

Query Match 0.5%; Score 7; DB 20; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 |||||  
 Db 215 slalvga 221

RESULT 99  
 AAB35131  
 ID AAB35131 standard; protein; 236 AA.

XX AAB35131;  
 XX 03-APR-2001 (first entry)  
 DT Murine Bcl-2.  
 DE Mouse; Bax; apoptosis modulator; BCL-2.  
 XX

OS Mus sp.  
 XX US6165732-A.  
 PN 26-DEC-2000.  
 PD 31-JUL-1998; 98US-0127048.  
 XX 14-OCT-1997; 97US-0061823.  
 PR (UNIV) UNIV WASHINGTON.  
 XX Korsmeyer SJ, Schlesinger PH;  
 FI WPI; 2001-101692/11.  
 XX Identifying apoptosis-modulating compounds by contacting the compound  
 PT with lipid bilayer containing an ion channel formed by anti-apoptotic  
 PT polypeptide of Bcl-2 family and determining ion selectivity of the  
 PT channel -  
 XX Example 1; Fig 11; 34pp; English.  
 XX The present invention describes a method for identifying modulators of  
 CC apoptosis which involves contacting a compound of interest with a lipid  
 CC bilayer comprising a K<sup>+</sup> or Cl<sup>-</sup> selective channel. This channel is a  
 CC member of the BCL-2 family. Apoptosis modulators are also provided,  
 CC including Bcl-2deltaTM and BaxdeltaTM.  
 XX Sequence 236 AA;

Query Match 0.5%; Score 7; DB 22; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 |||||  
 Db 219 slalvga 225

RESULT 100  
 AAB80987  
 ID AAB80987 standard; protein; 239 AA.

XX AAB80987;  
 XX 17-DEC-1990 (first entry)  
 DT Sequence of bcl-2-alpha encoded by sequence of bcl-2 cDNA corresp. to  
 DE the 5.5 kb transcript.  
 DE B-cell neoplasm; diagnosis; follicular lymphomas.

XX Homo sapiens.  
 XX EP252685-A.  
 PN 13-JAN-1988.  
 PD 02-JUL-1987; 87EP-0305863.  
 XX 09-JUL-1986; 86US-0883687.

XX (WIST-) WISTAR CORP.  
 XX Tsujimoto Y, Croce CM;  
 PI WPI; 1988-008633/02.  
 DR N-PSDB; AAN81292.  
 XX

PT Detection of B-cell neoplasms -  
 PT by extn. of proteins or RNA from B-cells and quantitation using

PT specific antibody or DNA probe

XX Claim 12; Fig 2A-2D; 23pp; English.

XX A human bcl-2 gene substantially free of introns is claimed. Also  
 CC claimed is a substantially pure preparation of a protein having an  
 CC N-terminal end encoded by the first exon of the human bcl-2 gene wherein  
 CC said protein is bcl-2-alpha having about 239 (AAP80987) or 205 (AAP80988)  
 CC AA residues. B-cell neoplasms which are associated with t(14;18)  
 CC chromosome translocations cause an increase in expression of both the  
 CC mRNA and the protein prods. of the bcl-2 gene. This is used to detect  
 CC B-cell neoplasms including follicular lymphomas as well as other  
 CC lymphomas. Bacterial isolates available as ATCC 67147 and 67148 can be  
 CC used to express bcl-2 gene products alpha (AAN81292) and beta (AAN81293)  
 CC resp. in bacteria.

XX Sequence 239 AA;

Query Match 0.5%; Score 7; DB 9; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22  
 |||||

Db 222 slalvga 228

RESULT 101

AAR42312  
 ID AAR42312 standard; Protein; 239 AA.

AC AAR42312;

DT 03-MAY-1994 (first entry)

DE Bcl-2 oncogene product.

KW Cell death; apoptosis; inhibition; de-inhibition; bcl-2 oncogene;  
 KW expression; myc.

OS Homo sapiens.

PN W09320200-A.

PD 14-OCT-1993.

PF 02-APR-1993; 93WO-GB00686.

PR 02-APR-1992; 92GB-0007275.

PR 02-APR-1992; 92GB-0007276.

XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.

XX Evan GI;

DR WPI: 1993-336908/42.  
 DR N-PSDB; AAQ49815.

XX Treating tumour cells by de-inhibiting Myc-induced apoptosis -  
 PT esp. by inhibiting expression of the BCL-2 oncogene e.g. with  
 PT antisense oligo:nucleotide(s), also increasing survival of  
 PT cultured cells by expressing BCL-2

XX Disclosure; Page 76-77; 109pp; English.

XX A DNA construct comprising the bcl-2 coding sequence under control  
 CC of elements allowing its expression is claimed. Myc-induced cell  
 CC death can be inhibited in cultured cells by expressing bcl-2.  
 CC Myc-induced cell death can be de-inhibited in tumour cells by admin.  
 CC of bcl-2 antisense oligonucleotides.

XX Sequence 239 AA;

Query Match 0.5%; Score 7; DB 14; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22  
 |||||

Db 222 slalvga 228

RESULT 102

AAR47344  
 ID AAR47344 standard; Protein; 239 AA.

XX AC AAR47344;

XX DT 23-JUN-1994 (first entry)

XX Human oncogene bcl-2 product.  
 DE

XX Cell death; senescence; programmed cell death; ced-9; myocardial  
 KW infarction; stroke; brain injury; neurodegenerative disease;  
 KW muscular degenerative disease; ageing; hypoxia; ischaemia; toxemia;  
 KW infection; hair loss; neoplasia; cancer; ced-3; ced-4; bcl-2;  
 KW oncogene.

XX Homo sapiens.

XX W09325683-A.

XX PD 23-DEC-1993.

XX PF 14-JUN-1993; 93WO-US05651.

PR 12-JUN-1992; 92US-0898933.

PR 10-AUG-1992; 92US-0927681.

XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.

XX Hengartner M, Horvitz HR;

DR WPI: 1994-007540/01.

DR N-PSDB; AAQ54631.

XX Caenorhabditis elegans cell death-protective gene - used to  
 PT develop agents for preventing cell death or for reducing  
 PT population of cells

XX Disclosure; Page 64-65; 112pp; English.

XX The protein product of the human oncogene bcl-2 was found to have a  
 CC similar sequence to the ced-9 protein. ced-9 is essential for  
 CC C. elegans development and apparently functions by protecting cells  
 CC during development from programmed cell death. ced-9 was shown to  
 CC function by antagonising the activities of cell death genes ced-3  
 CC and ced-4. The ced-9 gene can be used for developing agents for  
 CC treating a condition characterised by increased cell death such as  
 CC myocardial infarction, stroke, traumatic brain injury,  
 CC neurodegenerative disease, muscular degenerative disease, ageing,  
 CC hypoxia, ischaemia, toxemia, infection or hair loss. It can also  
 CC be used for reducing a population of cells in the treatment of  
 CC neoplastic growth cancerous tissue, infected cells or autoreactive  
 CC immune cells.

XX Sequence 239 AA;

Query Match 0.5%; Score 7; DB 15; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLALVGA 22

```

Db      222 slalvga 228
|||||||
RESULT 103
AAR70331
ID AAR70331 standard; Protein; 239 AA.
XX AC AAR70331;
XX XX
XX 27-SEP-1995 (first entry)
XX DE Human bcl-2 protein.
XX KW Anticodon oligomer; antisense oligonucleotide; bcl-2; cancer; therapy;
XX KW chemoresistance.
XX OS Homo sapiens.
XX PN W09508350-A.
XX XX
XX PD 30-MAR-1995.
XX PF 20-SEP-1994; 94WO-US10725.
XX PR 20-SEP-1993; 93US-0124256.
XX PA (REED/) REED J C.
XX PI Reed JC;
XX DR WPI; 1995-139394/18.
XX PT Anti-code oligomers which bind to bcl-2 mRNA - for the treatment
XX PT of human solid tumours, esp. breast cancer
XX PS Disclosure; Page 71-72; 108pp; English.
XX CC The human bcl-2 gene encodes a 25 kDa protein (AAR70331). Antisense
XX CC oligonucleotides have been designed to bind sites in mRNA transcribed
XX CC from the bcl-2 gene, thereby reducing expression of the bcl-2 protein
XX CC and inducing cell death in certain cancer cells.
XX SQ Sequence 239 AA;

Query Match 0.5%; Score 7; DB 16; Length 239;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22
|||||||
Db 222 slalvga 228

RESULT 104
AAR71404
ID AAR71404 standard; protein; 239 AA.
XX AC AAR71404;
XX XX
XX 30-OCT-1995 (first entry)
XX DE Human bcl-2 alpha protein.
XX KW Human; bcl-2; alpha; beta; proto-oncogene; hematopoietic cell line;
XX KW apoptosis; membrane-associated cytoplasmic protein; B cell; T cell;
XX KW proliferation; cell cycle progression; Bax; apoptotic cell death;
XX KW apoptosis; cytokine; death repressor; BHL; BH2; cancer therapy;
XX KW hyperplasia; immunodeficiency disease; AIDS; neurodegeneration;
XX KW ischaemic cell death.
XX OS Homo sapiens.

```

```

XX Key Location/Qualifiers
FH Domain 136..155
FT /label= BHL_domain
FT /note= "Represents Bax binding site"
FT 187..202
FT /label= BH2_domain
FT /note= "Represents Bax binding site"
XX
XX WO9505750-A.
PN 02-MAR-1995.
XX
XX 24-AUG-1994; 94WO-US09701.
XX 26-AUG-1993; 93US-0112208.
XX 25-MAY-1994; 94US-0248819.
XX (UNIW ) UNIV WASHINGTON.
XX
XX Korsmeyer SJ;
XX WPI; 1995-106605/14.
XX
XX Methods for producing and identifying mutant bcl-2 proteins -
XX that lack death repressor activity and/or lacks binding to Bax.
XX
XX Disclosure; Page 39; 133pp; English.
XX
XX The sequences given in AAR71404-05 represent the human bcl-2 alpha and
XX beta proteins respectively. bcl-2 is encoded by a proto-oncogene and
XX is capable of inhibiting apoptosis in many hematopoietic cell systems.
XX bcl-2 is a 26 kD membrane-associated cytoplasmic protein and is thought
XX to function by enhancing the survival of hematopoietic cells of B and T
XX origins rather than directly promoting proliferation of these cell
XX types. bcl-2 has not been shown to directly promote cell cycle
XX progression nor does it necessarily alter the dose response to limiting
XX concentrations of IL-3. bcl-2 has been shown to form heterodimers with
XX a 21 kD protein, Bax. Overexpressed Bax accelerates apoptotic cell
XX death induced by cytokine deprivation in an IL-3 dependent cell line, and
XX it also acts to counter the death repressor activity of bcl-2.
XX Therefore, the ratio between bcl-2 and Bax determines cell survival or
XX death following an apoptotic stimulus. The invention gives a mutant form
XX of bcl-2 in which there is at least one amino acid substitution or
XX deletion in the BHL or BH2 domains. This makes the mutant protein
XX substantially incapable of binding Bax and/or incapable of death
XX repressor activity. Down regulation of bcl-2 is useful in cancer therapy,
XX controlling hyperplasias and eliminating self-reactive clones in
XX autoimmunity by favouring death effector molecules. Up regulating
XX bcl-2 is beneficial in treatment and diagnosis of immunodeficiency
XX diseases, including AIDS and neurodegenerative and ischaemic cell death.
XX
XX SQ Sequence 239 AA;

Query Match 0.5%; Score 7; DB 16; Length 239;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22
|||||||
Db 222 slalvga 228

RESULT 105
AAW02383
ID AAW02383 standard; Protein; 239 AA.
XX AC AAW02383;
XX XX
XX 04-JUN-1997 (first entry)
XX DT
XX DE Human BCL2.

```

XX Sense oriented; genetic suppressor element; GSE; reverse;  
 KW BCL2; gene; mediated; suppression; apoptosis; mammalian; cell;  
 KW inhibition; sensitisation; cancer; chemotherapeutic agent;  
 KW increase; treatment; induction; virus; infection; death;  
 KW disease; haematopoietic; neurological; recombinant construct;  
 KW decrease; expression; anticancer; non-Hodgkin's lymphoma;  
 KW B cell malignancy.  
 XX Homo sapiens.  
 OS  
 XX  
 PN WO9629403-A1.  
 XX  
 PD 26-SEP-1996.  
 XX  
 XX 14-MAR-1996; 96WO-US03545.  
 XX  
 PF 17-MAR-1995; 95US-0405702.  
 PR  
 XX (UNII ) UNIV ILLINOIS FOUND.  
 PA  
 XX Holzmayr TA, Roninson IB, Schott B, Tarasiewicz DG;  
 PI  
 XX WPI; 1996-443179/44.  
 DR N-PSDB; AAT33694.  
 XX  
 XX Sense oriented genetic suppressor element - for reversing BCL2  
 PT mediated inhibition of apoptosis, and for sensitising cancer cells  
 PT against chemotherapeutic agents  
 PT  
 XX  
 PS Claim 20; Pages 37-38; 66pp; English.  
 XX  
 CC The present sequence is human BCL2 from which a peptide  
 CC capable of inhibiting BCL2 gene, or gene product, function in a  
 CC cell can be derived. The cDNA sequence encoding the peptide is a  
 CC sense oriented genetic suppressor element (GSE) for reversing  
 CC BCL2 mediated suppression of apoptosis in a mammalian cell. The  
 CC GSE and its peptide product can be used to sensitise cancer cells  
 CC to chemotherapeutic agents, and to increase apoptosis, especially  
 CC for the treatment of cancer, but more generally to induce virus  
 CC infected cell death, or to treat apoptosis related diseases of  
 CC haematopoietic or neurological cells. The GSE peptide product or a  
 CC recombinant construct encoding the GSE can be used to decrease  
 CC BCL2 gene expression by exerting an anticancer effect, e.g. in  
 CC cases of non-Hodgkin's lymphoma and B cell malignancy.  
 XX  
 SQ Sequence 239 AA;

Query Match 0.5%; Score 7; DB 17; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 |||||  
 Db 222 slalvga 228

RESULT 106  
 AAW01018  
 ID AAW01018 standard; Protein; 239 AA.  
 XX  
 AC AAW01018;  
 XX  
 DT 18-DEC-1996 (first entry)  
 XX  
 DE Apoptosis-blocking protein Bcl-2.  
 XX  
 KW Apoptosis-regulating protein; Bcl-2; oncogene;  
 KW adenovirus E1B 19k protein; cell death; cancer; tumour;  
 KW immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; Nip1;  
 KW Nip2; Nip3.  
 XX

OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Binding-site 43..51  
 FT /label= "Binding\_motif  
 FT /note= "Interacts with Bip proteins"  
 FT Binding-site 106..112  
 FT /label= "Binding\_motif  
 FT /note= "Interacts with Bip proteins"  
 XX  
 PN EP733706-A2.  
 XX  
 PD 25-SEP-1996.  
 XX  
 XX 21-MAR-1996; 96EP-0104542.  
 XX  
 PF 21-MAR-1995; 95US-0408095.  
 PR  
 XX (UYSL-) UNIV ST LOUIS.  
 PA  
 XX Chinnadurai G;  
 PI  
 XX WPI; 1996-427055/43.  
 DR  
 XX Nucleic acids encoding apoptosis regulating proteins - useful for  
 PT diagnosing and treating immune disorders, malignancies, etc.  
 PT  
 XX Example 8; Page 32-33; 60pp; English.  
 XX  
 CC The bcl-2 oncogene product (AAW01018) enhances the survival of  
 CC haematopoietic B and T cells by blocking apoptosis induced by  
 CC diverse agents. Its activity is similar to that of the 19K  
 CC protein (AAW01010) of adenovirus E1B. 3 Novel proteins, Bip1A,  
 CC Bip13 and Bip5 (AAW01000-02), that specifically interact with  
 CC Bcl-2, have been identified. Mutational analysis (see also  
 CC AAW01019-21) shows the apoptosis-regulating Nip proteins (see also  
 CC AAW00997-99) associate with Bcl-2 at specific sites (see also  
 CC AAW01003-04) that show homology to motifs (AAW01005-06) on 19K.  
 XX  
 SQ Sequence 239 AA;

Query Match 0.5%; Score 7; DB 17; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 |||||  
 Db 222 slalvga 228

RESULT 107  
 AAW40217  
 ID AAW40217 standard; Peptide; 239 AA.  
 XX  
 AC AAW40217;  
 XX  
 DT 07-JUL-1998 (first entry)  
 XX  
 DE Human bcl-2.  
 XX  
 KW TMAH; apoptosis; osteoarthritis; c-type lectin; A1 family; diagnosis;  
 KW treatment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9804585-A2.  
 XX  
 PD 05-FEB-1998.  
 XX  
 PF 22-JUL-1997; 97WO-US13077.  
 XX  
 PR 31-JUL-1996; 96US-0690095.

XX BA (INCY-) INCYTE PHARM INC.  
 XX PI Au-Young J, Goli SK, Hillman JL;  
 XX DR WPI; 1998-130617/12.  
 XX PT Human macrophage antigen - used for decreasing apoptosis associated  
 XX PT with osteoarthritis  
 XX PS Disclosure; Page 43; 58pp; English.  
 XX CC The human bcl-2 peptide is one of a group of peptides with which the  
 CC human macrophage antigen (TMAH)(AAW40215) has 20% homology. The  
 CC homology which TMAH shares with the other A1 family members includes  
 CC conserved residues at F27, P35, R119, W139, F146 AND W214. The  
 CC structural homology between the mammalian A1 and C-type lectins and TMAH  
 CC provides information on the structural and physical properties of both  
 CC the TMAH gene and protein. This is used in the development of TMAH as a  
 CC diagnostic tool and as a method of treating diseases associated with  
 CC expression of TMAH.  
 XX SQ Sequence 239 AA;

Query Match 0.5%; Score 7; DB 19; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 SLALVGA 22  
 | | | | |  
 DB 222 slalvga 228

RESULT 108  
 AAW94345  
 ID AAW94345 standard; Protein; 239 AA.  
 AC AAW94345;  
 XX DT 13-APR-1999 (first entry)  
 XX DE Human Bcl-2 wild-type protein.  
 XX KW Human; Nip1; Nip2; Nip3; Bip1a; Bip3; adenovirus; cell death;  
 KW viral infection; Bcl-2; protooncogene; mutational analysis; apoptosis;  
 KW ELB 19K protein; cell survival regulation.  
 XX OS Homo sapiens.  
 XX PN US5858678-A.  
 XX PD 12-JAN-1999.  
 XX PF 21-MAR-1995; 95US-0408095.  
 XX PR 02-AUG-1994; 94US-0284139.  
 XX PR 21-MAR-1995; 95US-0408095.  
 XX PA (UYSL-) UNIV ST LOUIS.  
 XX PI Chinnadurai G;  
 XX DR WPI; 1999-152099/13.  
 XX PT Polypeptides that bind to anti-apoptotic proteins - useful for  
 PT protecting against cell death induced by viral infection and to  
 PT modulate response to physical and chemical stimuli  
 XX PS Example 8; Column 41-42; 41pp; English.  
 XX CC The present invention describes: (1) a method for regulating cell death,  
 CC comprising exposing an isolated cell to a polypeptide selected from

CC Nip1, Nip2, Nip3, Bip1a, Bip5 and Bip13; (2) a method for neutralising  
 CC the activity of the adenovirus E1B 19 kD protein, the Bcl-2 protein or  
 CC the BHRF-1 protein, comprising exposing an isolated cell to a  
 CC polypeptide as in (1); and (3) a method for detecting molecules that  
 CC bind to at least one polypeptide as in (1), comprising lysing cells,  
 CC exposing the lysate to the polypeptide and detecting any molecule-  
 CC polypeptide aggregates. The methods are useful for providing proteins  
 CC able to bind to other proteins known to regulate cell survival e.g. it  
 CC is known that E1B 19K protein provides a survival function similar to  
 CC the cellular protooncogene bcl-2 gene product which is able to block  
 CC apoptosis in haematopoietic B and T cells. The present sequence  
 CC represents the human Bcl-2 wild-type protein from the present  
 CC invention.  
 XX SQ Sequence 239 AA;

Query Match 0.5%; Score 7; DB 20; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 SLALVGA 22  
 | | | | |  
 DB 222 slalvga 228

RESULT 109  
 AAW87810  
 ID AAW87810 standard; Protein; 239 AA.  
 XX AC AAW87810;  
 XX DT 10-MAR-1999 (first entry)  
 XX DE A human Bcl-2 protein.  
 XX KW Human; Bcl-2 associated protein; Bax; bcl-2; antibody; modulator;  
 KW bcl-2-related function; apoptosis; dimer; Bcl-xL; Mcl-1; A1.  
 XX OS Homo sapiens.  
 XX PN US5856171-A.  
 XX PD 05-JAN-1999.  
 XX PF 10-NOV-1994; 94US-0337646.  
 XX PR 10-NOV-1994; 94US-0337646.  
 XX PR 26-AUG-1993; 93US-0112208.  
 XX PR 25-MAY-1994; 94US-0248819.  
 XX PA (UNIW ) UNIV WASHINGTON.  
 XX PI Korsmeyer SJ;  
 XX DR WPI; 1999-105119/09.  
 XX PT DNA composition encoding bcl-2 two-hybrid and reporter system - for  
 PT identifying modulators of bcl-2 function.  
 XX PS Claim 6; Fig 7; 105pp; English.  
 XX CC The present sequence represents a human Bcl-2 protein. The  
 CC specification also describes Bcl-2 associated proteins  
 CC designated Bax. The Bax protein is used in a composition which  
 CC comprises a bcl-2 family member polypeptide, a naturally occurring  
 CC Bax polypeptide and an antibody that binds to the Bax polypeptide.  
 CC The specification also describes a composition comprising a hybrid  
 CC protein comprising an activator domain of a transcriptional activator  
 CC protein and a bcl-2 family member having a BHI domain and a BH2 domain;  
 CC another hybrid protein comprising a DNA-binding domain of the  
 CC transcriptional activator protein and a second bcl-2 family member  
 CC having a BHI domain and a BH2 domain; and a reporter gene linked to a

transcriptional regulatory element whose transcriptional activity is dependent on the presence or absence of a dimer of the two hybrid proteins. The bcl-2 family members are selected from naturally occurring Bcl-2, Bcl-xL, Bax, Mcl-1, A1, fragments thereof, and mutants having a mutation in the BH1 and/or BH2 domain that alters intermolecular binding of the two bcl-2 family members. The compositions are used to identify modulators of bcl-2-related function, e.g. substances that inhibit binding of Bax to bcl-2, which would be potentially useful as drugs for modulating apoptosis.

Sequence 239 AA;

Query Match 0.5%; Score 7; DB 20; Length 239;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 222 slalvga 228  
|||||

RESULT 110  
AAW87812  
ID AAW87812 standard; Protein; 239 AA.

AC AAW87812;

DT 10-MAR-1999 (first entry)

DE A human Bcl-2-alpha protein.

KW Human; Bcl-2 associated protein; Bax; bcl-2; antibody; modulator;  
KW bcl-2-related function; apoptosis; dimer; Bcl-xL; Mcl-1; A1.

OS Homo sapiens.

PN US5856171-A.

PD 05-JAN-1999.

PF 10-NOV-1994; 94US-0337646.

PR 10-NOV-1994; 94US-0337646.

PR 26-AUG-1993; 93US-0112208.

PR 25-MAY-1994; 94US-0248819.

PA (UNIW ) UNIV WASHINGTON.

PI Korsmeyer SJ;

DR WPI; 1999-105119/09.

PT DNA composition encoding bcl-2 two-hybrid and reporter system - for  
PT identifying modulators of bcl-2 function

PS Disclosure; Column 29; 105pp; English.

The present sequence represents a human Bcl-2-alpha protein. The specification also describes Bcl-2 associated proteins designated Bax. The Bax protein is used in a composition which comprises a bcl-2 family member polypeptide, a naturally occurring Bax polypeptide and an antibody that binds to the Bax polypeptide. The specification also describes a composition comprising a hybrid protein comprising an activator domain of a transcriptional activator protein and a bcl-2 family member having a BH1 domain and a BH2 domain; another hybrid protein comprising a DNA-binding domain of the transcriptional activator protein and a second bcl-2 family member having a BH1 domain and a BH2 domain; and a reporter gene linked to a transcriptional regulatory element whose transcriptional activity is dependent on the presence or absence of a dimer of the two hybrid proteins. The bcl-2 family members are selected from naturally occurring Bcl-2, Bcl-xL, Bax, Mcl-1, A1, fragments thereof, and mutants having a

CC mutation in the BH1 and/or BH2 domain that alters intermolecular binding  
CC of the two bcl-2 family members. The compositions are used to identify  
CC modulators of bcl-2-related function, e.g. substances that inhibit  
CC binding of Bax to bcl-2, which would be potentially useful as drugs  
CC for modulating apoptosis.

XX Sequence 239 AA;

Query Match 0.5%; Score 7; DB 20; Length 239;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
Db 222 slalvga 228  
|||||

RESULT 111  
AAY69203  
ID AAY69203 standard; peptide; 239 AA.

XX AC AAY69203;

DT 30-MAY-2000 (first entry)

DE Amino acid sequence of a human Bcl-2 protein.

KW Pro-apoptotic peptide; channel inducer; transport;

KW cytochrome C transport; mitochondria; apoptosis; ion selectivity;

KW anti-apoptotic Bcl-2 family member; neoplasia; Epstein Barr virus;

KW African swine fever virus; adenovirus; lymphoproliferative condition;

KW cancer; arthritis; Crohn's disease; inflammation; autoimmune disease;

KW immunodeficiency; senescence; neurodegenerative disease;

KW reperfusion cell death; infertility; wound; Bcl-2.

XX Homo sapiens.

XX WO200006187-A2.

XX 10-FEB-2000.

XX 30-JUL-1999; 99WO-US17276.

XX 31-JUL-1998; 98US-0127048.

XX (UNIW ) UNIV WASHINGTON.

XX Korsmeyer SJ, Schlesinger PH;

XX WPI; 2000-195193/17.

XX Modulating apoptosis in cells by modulating channel ion selectivity for  
XX transport of cytochrome C -

XX Disclosure; Page 34; 57pp; English.

The present sequence represents the Bcl-2 protein, from which an anti-apoptotic peptide is derived. The peptide is an inhibitor of formation of a channel for transport of cytochrome C out of mitochondria. The peptide inhibits apoptosis in a cell. The peptide changes the ion selectivity of an anti-apoptotic Bcl-2 family member from chloride selective to potassium selective. The specification also describes inducers of apoptosis in cells. The inhibitors and inducers can be used to treat patients, preferably humans with a condition mediated by excessive down-regulation of apoptosis, especially conditions chosen from neoplasias, diseases caused by Epstein Barr virus, African swine fever virus and adenovirus, lymphoproliferative conditions, cancer, arthritis, Crohn's disease, inflammation and autoimmune disease or a condition mediated by excessive apoptosis, especially immunodeficiency diseases, senescence, neurodegenerative disease, ischemic and reperfusion cell death, infertility and wounds. The methods can also be used to identify apoptosis-modulating compounds.



XX  
SQ

Sequence 239 AA;

Query Match 0.5%; Score 7; DB 21; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 Db 222 slalvga 228  
 |||||

RESULT 112

AAB74127  
 ID AAB74127 standard; Protein; 239 AA.

XX AC AAB74127;

XX DT 22-MAY-2001 (first entry)  
 XX DE Human bcl-2.

XX KW Human; Bax; cytostatic; immunosuppressive; immunostimulant; infection;  
 KW apoptosis modulator; bcl-2 associated X protein; cancer therapy; AIDS;  
 KW autoimmunity; immunodeficiency; reperfusion injury; stroke; aging;  
 KW myocardial infarction; traumatic brain injury; ischaemia; bcl-2;  
 KW neurodegenerative diseases; hepatitis; transplant rejection; toxemia;  
 KW lymphoproliferative disease; chromosome 18q21.3.

XX OS Homo sapiens.  
 XX PN US6184202-B1.

XX PD 06-FEB-2001.  
 XX PF 11-SEP-1997; 97US-0927326.

XX PR 10-NOV-1994; 94US-0337646.  
 XX PR 26-AUG-1993; 93US-0112208.  
 XX PR 25-MAY-1994; 94US-0248819.

XX PA (UNIW ) UNIV WASHINGTON.  
 XX PI Korsmeyer SJ;

XX DR WPI; 2001-256104/26.  
 XX PT Modulating apoptosis of a cell, useful in maintaining homeostasis in

XX PT adult tissues, or treating proliferative or autoimmune diseases,  
 XX PT comprises administering a bcl-2 polypeptide that interacts with a 21 kD  
 XX PT bcl-2 associated X protein -

XX PS Claim 3; Fig 7; 105pp; English.

XX CC The present invention relates to a method of modulating apoptosis of a  
 CC cell. The method comprises administering to the cell an agent,  
 CC comprising a BH1 domain or BH2 domain, capable of modulating formation of  
 CC at least one complex selected from bcl-2:bcl-2 complexes, bcl-XL:bcl-XL  
 CC complexes, bcl-2 associated X protein (Bax):Bax complexes, bcl-2:Bax  
 CC complexes or bcl-XL:Bax complexes. Modulating apoptosis is especially  
 CC useful in cancer therapy, and treating autoimmunity, immunodeficiency  
 CC diseases (e.g. AIDS), reperfusion injury, myocardial infarction, stroke,  
 CC traumatic brain injury, neurodegenerative diseases, aging, ischaemia,  
 CC toxemia, infection, hepatitis, transplant rejection, and  
 CC lymphoproliferative diseases. The present sequence is human Bcl-2, which  
 CC was used in a sequence alignment in the present invention, with murine  
 CC Bax (AAB74125), human Bax (AAB74126) and murine Bcl-2 (AAB74128). The  
 CC human Bcl-2 gene is located on chromosome 18q21.3.

XX SQ Sequence 239 AA;

Query Match 0.5%; Score 7; DB 22; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 Db 222 slalvga 228  
 |||||

RESULT 113

AAB74129  
 ID AAB74129 standard; Protein; 239 AA.

XX AC AAB74129;

XX DT 22-MAY-2001 (first entry)  
 XX DE Human bcl-2alpha.

XX KW Human; Bax; cytostatic; immunosuppressive; immunostimulant; infection;  
 KW apoptosis modulator; bcl-2 associated X protein; cancer therapy; AIDS;  
 KW autoimmunity; immunodeficiency; reperfusion injury; stroke; aging;  
 KW myocardial infarction; traumatic brain injury; ischaemia; bcl-2alpha;  
 KW neurodegenerative diseases; hepatitis; transplant rejection; toxemia;  
 KW lymphoproliferative disease; chromosome 18q21.3.

XX OS Homo sapiens.  
 XX PN US6184202-B1.

XX PD 06-FEB-2001.  
 XX PF 11-SEP-1997; 97US-0927326.

XX PR 10-NOV-1994; 94US-0337646.  
 XX PR 26-AUG-1993; 93US-0112208.  
 XX PR 25-MAY-1994; 94US-0248819.

XX PA (UNIW ) UNIV WASHINGTON.  
 XX PI Korsmeyer SJ;

XX DR WPI; 2001-256104/26.  
 XX PT Modulating apoptosis of a cell, useful in maintaining homeostasis in

XX PT adult tissues, or treating proliferative or autoimmune diseases,  
 XX PT comprises administering a bcl-2 polypeptide that interacts with a 21 kD  
 XX PT bcl-2 associated X protein -

XX PS Disclosure; Columns 29-30; 105pp; English.

XX CC The present invention relates to a method of modulating apoptosis of a  
 CC cell. The method comprises administering to the cell an agent,  
 CC comprising a BH1 domain or BH2 domain, capable of modulating formation of  
 CC at least one complex selected from bcl-2:bcl-2 complexes, bcl-XL:bcl-XL  
 CC complexes, bcl-2 associated X protein (Bax):Bax complexes, bcl-2:Bax  
 CC complexes or bcl-XL:Bax complexes. Modulating apoptosis is especially  
 CC useful in cancer therapy, and treating autoimmunity, immunodeficiency  
 CC diseases (e.g. AIDS), reperfusion injury, myocardial infarction, stroke,  
 CC traumatic brain injury, neurodegenerative diseases, aging, ischaemia,  
 CC toxemia, infection, hepatitis, transplant rejection, and  
 CC lymphoproliferative diseases. The present sequence is human Bcl-2alpha,  
 CC which was used in the method of the present invention. The human Bcl-2  
 CC gene is located on chromosome 18q21.3.

XX SQ Sequence 239 AA;

Query Match 0.5%; Score 7; DB 22; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22

Db 222 slalvga 228  
|||||

## RESULT 114

AAB48288  
ID AAB48288 standard; protein; 239 AA.

XX AC AAB48288;

DT 02-APR-2001 (first entry)

XX DE Human BCL-2 protein.

XX KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; 2F;  
KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;  
KW Bad; Bcl-2; tumour; cytosolic.

XX OS Homo sapiens.

XX PN WO200075184-A1.

XX PD 14-DEC-2000.

XX 05-JUN-2000; 2000WO-US15449.

XX 04-JUN-1999; 99US-0137494.

XX (UYA ) UNIV YALE.

PI Zhang H, Tsvetkov LM, Kondo T;

DR WPI: 2001-061703/07.

DR N-PSDB; AAC84600.

XX Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
PT involves altering levels of proteins such as S-phase kinase associated  
PT proteins 1, 2 and cullin/CDC53 proteins -

PS Claim 5; Page 104-108; 162pp; English.

XX The invention relates to methods of altering the polypeptide levels in a  
CC cell, using proteins selected from S-phase kinase associated proteins 1  
CC and 2 (SKP1, SKP2), SKP2-like proteins (2F) and CUL-1 (a member of the  
CC cullin/CDC53 family of proteins). The method is useful for altering the  
CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2  
CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for  
CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
CC that modulate interactions between SKP and target proteins are useful for  
CC treating tumours.

XX Sequence 239 AA;

Query Match 0.5%; Score 7; DB 22; Length 239;

Best Local Similarity 100.0%; Pred. No. 2.4e-02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22

|||||

Db 222 slalvga 228

## RESULT 115

AAB35130

ID AAB35130 standard; protein; 239 AA.

XX AC AAB35130;

DT 03-APR-2001 (first entry)

XX DE Human Bcl-2.

KW Human; Bax; apoptosis modulator; BCL-2.

XX OS Homo sapiens.

XX PN US6165732-A.

XX PD 26-DEC-2000.

XX 31-JUL-1998; 98US-0127048.

XX 14-OCT-1997; 97US-0061823.

XX (UNIW ) UNIV WASHINGTON.

XX Kormeyer SJ, Schlesinger PH;

XX WPI: 2001-101692/11.

XX Identifying apoptosis-modulating compounds by contacting the compound  
PT with lipid bilayer containing an ion channel formed by anti-apoptotic  
PT polypeptide of Bcl-2 family and determining ion selectivity of the  
PT channel -

XX Disclosure; Fig 11; 34pp; English.

XX The present invention describes a method for identifying modulators of  
CC apoptosis which involves contacting a compound of interest with a lipid  
CC bilayer comprising a K+ or Cl- selective channel. This channel is a  
CC member of the BCL-2 family. Apoptosis modulators are also provided,  
CC including Bcl-2deltaTM and BaxdeltaTM.

XX Sequence 239 AA;

Query Match 0.5%; Score 7; DB 22; Length 239;

Best Local Similarity 100.0%; Pred. No. 2.4e-02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22

|||||

Db 222 slalvga 228

## RESULT 116

AAB50537

ID AAB50537 standard; Protein; 239 AA.

XX AC AAB50537;

XX 16-MAR-2001 (first entry)

XX Human Bcl-2 protein sequence SEQ ID NO:2.

XX Human; Bcl-2; Bcl-xL; Bax; VDAC; apoptosis inhibitor; detection;

XX apoptosis promoter; diagnosis.

XX OS Homo sapiens.

XX JP2000287689-A.

XX 17-OCT-2000.

XX 08-APR-1999; 99JP-0101888.

XX 08-APR-1999; 99JP-0101888.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI: 2001-065575/08.

XX N-PSDB; AAC90809.

XX Screening of an apoptosis inhibitor or promoter which can be used as a  
PT drug and a diagnostic agent for various diseases caused by apoptosis

XX inhibition or apoptosis promotion -  
 PS Claim 11; Page 13-14; 22pp; Japanese.  
 XX The present invention describes a method for screening for an apoptosis  
 CC inhibitor or an apoptosis promoter in which VDAC-liposome, an index  
 CC substance which can pass VDAC and a sample are incubated and the change  
 CC in the concentration of the index substance during the incubation is  
 CC detected to judge the presence of apoptosis inhibition or apoptosis  
 CC promotion. The apoptosis inhibitor or the apoptosis promoter can be  
 CC used as a drug and a diagnostic agent for various diseases caused by  
 CC apoptosis inhibition or apoptosis promotion. The present sequence  
 CC represents the human Bcl-2 protein, which is an apoptosis inhibitor  
 CC used in the exemplification of the present invention.  
 XX  
 SQ Sequence 239 AA;  
 Query Match 0.5%; Score 7; DB 22; Length 239;  
 Best Local Similarity 100.0%; Pred. NO. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 16 SLALVGA 22  
 Db 222 slalvga 228  
 |||||  
 RESULT 117  
 AAW46943  
 ID AAW46943 standard; Protein; 245 AA.  
 XX  
 AC AAW46943;  
 XX  
 DT 17-AUG-1998 (first entry)  
 DE DR-alpha extracellular domain-Fos leucine zipper fusion protein.  
 XX  
 KW Major histocompatibility complex class II; MHC class II; human;  
 KW fusion protein; HLA-DR2; DRA\*0101; binding domain; Fos;  
 KW dimerisation domain; allergy; autoimmune disease; vaccine;  
 KW multiple sclerosis; therapy.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..7  
 FT /label= Sig\_peptide  
 FT Protein 8..245  
 FT /label= Mat\_protein  
 FT Domain 8..198  
 FT /note= "DR-alpha extracellular domain"  
 FT Peptide 199..205  
 FT /note= "linker"  
 FT Domain 206..245  
 FT /note= "Fos leucine zipper dimerisation domain"  
 FT Cleavage-site 6..7  
 FT /note= "KEX2 cleavage site"  
 XX  
 PN WO9806749-A2.  
 XX  
 PD 19-FEB-1998.  
 XX  
 PF 15-AUG-1997; 97WO-US14503.  
 XX  
 PR 16-AUG-1996; 96US-0024077.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Strominger JL, Wucherpennig KW;  
 XX  
 DR WPI; 1998-159459/14.  
 DR N-PSDB; AAV16866.

XX New Class II MHC fusion proteins - comprising a MHC Class II  
 PT binding domain and a dimerisation domain or an immunoglobulin region  
 PT used for modulating immune responses  
 XX  
 PS Example; Page 55; 76pp; English.  
 XX  
 CC This polypeptide comprises a fusion protein composed of the  
 CC extracellular domain of DR-alpha (residues 1-191 of DRA\*0101) and  
 CC the leucine zipper dimerisation domain of Fos, linked by a  
 CC 7-amino acid linker. A nucleotide construct (see AAV16866) encoding  
 CC the fusion was cloned into pPIC9, and the DR-alpha-Fos protein was  
 CC expressed in Pichia pastoris host cells. A DR-beta-Jun fusion (see  
 CC AAW46944) was also produced. CHO transfectants assembled and  
 CC secreted DR alpha-beta heterodimers. The DR2 molecules were shown  
 CC to specifically bind a human myelin basic protein that is  
 CC recognised by DR2-restricted T cell clones from multiple sclerosis  
 CC patients. The invention relates to new soluble monovalent or  
 CC multivalent Class II MHC fusion proteins comprising a MHC Class II  
 CC binding domain and a dimerisation domain or an immunoglobulin  
 CC region that can be used for the treatment of allergic and  
 CC autoimmune diseases, for tolerising a subject to foreign tissue  
 CC before or after organ or tissue transplantation, or for vaccination  
 CC against pathogens.  
 XX  
 SQ Sequence 245 AA;  
 Query Match 0.5%; Score 7; DB 19; Length 245;  
 Best Local Similarity 100.0%; Pred. NO. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1109 FEAQGAL 1115  
 Db 61 feaqgal 67  
 |||||  
 RESULT 118  
 AAY31652  
 ID AAY31652 standard; Protein; 245 AA.  
 XX  
 AC AAY31652;  
 XX  
 DT 09-NOV-1999 (first entry)  
 DE HLA-DR2 alpha-Fos leucine zipper fusion protein.  
 XX  
 KW Major histocompatibility complex Class II; MHC; binding domain;  
 KW HLA-DR2; leucine zipper; Fos; fusion protein; multiple sclerosis;  
 KW rheumatoid arthritis; graft rejection; allergy; autoimmune disease;  
 KW pemphigus vulgaris; systemic lupus erythematosus; T lymphocyte;  
 KW T cell; diagnosis; therapy; adoptive immunotherapy.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Saccharomyces cerevisiae.  
 OS Chimeric - synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..7  
 FT /note= "alpha-mating factor secretion signal"  
 FT Protein 8..245  
 FT /note= "mature protein"  
 FT Domain 8..198  
 FT /note= "DRA\*0101 extracellular domain"  
 FT Peptide 199..205  
 FT /note= "linker"  
 FT Domain 206..245  
 FT /note= "Fos leucine zipper domain"  
 XX  
 PN WO9942597-A1.  
 XX  
 PD 26-AUG-1999.  
 XX

PF 19-FEB-1999; 99WO-US03603.  
 XX  
 PR 19-FEB-1998; 98US-0075351.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Strominger JL, Wuchterpfennig KW;  
 XX  
 DR WPI; 1999-527481/44.  
 DR N-PSDB; AAX87807.  
 XX  
 PT New MHC Class II binding domain fusion proteins and conjugates -  
 PT used for, e.g. treating allergic and autoimmune diseases or  
 PT detecting, isolating, activating or killing specific T cells  
 XX  
 PS Example 1; Page 96-97; 113pp; English.  
 XX  
 CC The present sequence represents a fusion protein comprising an  
 CC alpha-mating factor secretion signal, the extracellular domain of  
 CC the HLA-DR2 alpha chain (residues 1-191 of DR\*0101), a 7-amino  
 CC acid linker, and the 40-amino acid leucine zipper dimerization  
 CC domain of Fos. The construct was expressed in Pichia pastoris  
 CC transformed host cells (see also AAX87807). The invention  
 CC provides new monovalent, multivalent and multimeric MHC Class II  
 CC binding domain fusion proteins and conjugates comprising at least  
 CC an MHC Class II binding domain of an MHC Class II alpha or beta  
 CC chain and a dimerization domain, especially a Fos or Jun leucine  
 CC zipper domain. The MHC fusion proteins and conjugates can be used  
 CC for detecting and isolating T cells having a defined MHC/peptide  
 CC complex specificity (claimed). They can also be used: to confer to  
 CC a subject adoptive immunity to a defined MHC/peptide complex  
 CC (claimed); to stimulate or activate T cells reactive to a defined  
 CC MHC/peptide complex (claimed); for selective killing of T cells  
 CC reactive to a defined MHC complex (claimed); to tolerate a subject  
 CC to a defined MHC/peptide complex (claimed); to treat allergic and  
 CC autoimmune diseases, e.g. multiple sclerosis, rheumatoid arthritis,  
 CC pemphigus vulgaris, and systemic lupus erythematosus; and to  
 CC prevent organ or tissue transplant rejection.  
 XX  
 SQ Sequence 245 AA;

Query Match 0.5%; Score 7; DB 20; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAGAL 1115  
 Db 61 feaqgal 67  
 |||||

RESULT 119  
 AAW37341  
 ID AAW37341 standard; Protein; 248 AA.  
 XX  
 AC AAW37341;  
 XX  
 DT 11-MAY-1998 (first entry)  
 XX  
 DE DR alpha-DAF chimeric protein.  
 XX  
 KW Vaccine; B-cell malignancy; lymphoma; leukaemia; tumour;  
 KW gene amplification; immunotherapy; therapy; DAF;  
 KW decay accelerating factor; DR alpha; mouse.  
 XX  
 OS Chimeric - Mus musculus.  
 XX  
 PN WO9741244-A1.  
 XX  
 PD 06-NOV-1997.  
 XX  
 PF 25-APR-1997; 97WO-US07039.  
 XX

PR 06-DEC-1996; 96US-0761277.  
 PR 01-MAY-1996; 96US-0644664.  
 XX  
 PA (GENI-) GENITOPE CORP.  
 XX  
 PI Denney DW;  
 XX  
 DR WPI; 1997-549743/50.  
 DR N-PSDB; AAT97175.  
 XX  
 PT Multivalent vaccine to treat B cell lymphoma or leukaemia -  
 PT comprises at least 2 different recombinant variable regions of  
 PT immunoglobulin molecules derived from B cell lymphoma cells  
 XX  
 PS Example 5; Page 114-115; 177pp; English.  
 XX  
 CC This sequence comprises a chimeric DR alpha-DAF protein in which  
 CC the extracellular domain of mouse DR-alpha is joined to sequences  
 CC derived from decay accelerating factor (DAF). The DAF sequence  
 CC provides a glycosylphosphatidylinositol linkage which allows the  
 CC chimeric protein to be cleaved from the surface of a cell by  
 CC treatment of the cell with phospholipase C. The chimeric gene  
 CC (see AAT97175) encoding the chimeric protein was used in the  
 CC construction of vector pSR alpha SD5-DR alpha-DAF. The invention  
 CC provides vectors and improved methods for the expression and  
 CC co-amplification of genes encoding recombinant proteins in  
 CC cultured cells. The amplified cells provide large quantities of  
 CC recombinant proteins suitable for immunotherapy for treatment of  
 CC lymphomas and leukaemia. The methods permit the production of  
 CC custom vaccines, including multivalent vaccines, that reflect  
 CC the degree of somatic variation found in a patient's tumour.  
 XX  
 SQ Sequence 248 AA;

Query Match 0.5%; Score 7; DB 18; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAGAL 1115  
 Db 79 feaqgal 85  
 |||||

RESULT 120  
 AAW98711  
 ID AAW98711 standard; Protein; 251 AA.  
 XX  
 AC AAW98711;  
 XX  
 DT 31-MAR-1999 (first entry)  
 XX  
 DE H. pylori GHPO 771 protein.  
 XX  
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9843478-A1.  
 XX  
 PD 08-OCT-1998.  
 XX  
 PF 01-APR-1998; 98WO-US06371.  
 XX  
 PR 29-JUL-1997; 97US-0902615.  
 PR 01-APR-1997; 97US-0833457.  
 PR 24-JUN-1997; 97US-0881227.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 XX  
 PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;

XX WPI; 1998-542293/46.  
 DR N-PSDB; AAX14430.  
 XX  
 PT New isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 PT infections and gastrointestinal diseases  
 XX  
 PS Claim 8; Page 1500-1501; 2054pp; English.  
 XX  
 CC This sequence represents a Helicobacter pylori GHPO protein of the  
 CC invention. The polypeptides can be used for preventing or treating  
 CC Helicobacter infections, and gastrointestinal diseases associated with  
 CC these infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
 CC used for the production of antibodies. The products can also be used for  
 CC detection and diagnosis.  
 XX  
 SQ Sequence 251 AA;

Query Match 0.5%; Score 7; DB 19; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1260 ELKLAKE 1266  
 Db 175 elklake 181  
 |||||

RESULT 121  
 AAY68277  
 ID AAY68277 standard; Peptide; 253 AA.  
 XX AC AAY68277;  
 XX  
 DT 13-APR-2000 (first entry)  
 XX  
 DE Class II alpha chain protein DR-alpha SEQ ID NO:109.  
 XX  
 KW MHC class I; major histocompatibility complex; microglobulin; antigen;  
 KW immune response; immunisation; AIDS; multiple sclerosis; toxic shock;  
 KW cancer; lupus erythematosus; snake bite; cytostatic; antiviral;  
 KW immunomodulatory; dermatological; immunosuppressive; antiinflammatory;  
 KW neuroprotective.  
 XX  
 OS Unidentified.  
 XX  
 PN US6011146-A.  
 XX  
 PD 04-JAN-2000.  
 XX  
 PF 07-JUN-1995; 95US-0481985.  
 XX  
 PR 15-NOV-1991; 91US-0792473.  
 PR 05-DEC-1991; 91US-0801818.  
 XX  
 PA (INSP ) INST PASTEUR.  
 PA (INRM ) INST NAT SANTE & RECH MEDICALE.  
 XX  
 PI Kourilsky P, Mottez E, Abastado J;  
 XX  
 DR WPI; 2000-125951/11.  
 XX  
 PT New recombinant DNA encoding covalently linked form of major  
 PT histocompatibility complex Class I determinant, used for immune system  
 PT stimulation, e.g. for treating cancer  
 XX  
 PS Disclosure; Column 129-132; 88pp; English.  
 XX  
 CC The present invention describes a recombinant DNA molecule (I)  
 CC containing a sequence (Ia) that encodes an altered MHC (major  
 CC histocompatibility complex ) Class I determinant (II) comprises a

CC polypeptide with alpha1, alpha2, alpha3 and beta2-microglobulin  
 CC domains, in which alpha3 and beta2 are covalently linked, thorough C-  
 CC and N-termini respectively, via a nucleotide spacer sequence encoding a  
 CC polypeptide. (II) includes an antigen-binding site and when (II) and  
 CC the antigen are associated they are recognized by a mammalian T cell  
 CC receptor (TCR). (I) are used to produce (II) which are used to study  
 CC functional interactions between the various MHC domains. They can also  
 CC be used to modulate (in vivo or in vitro) the immune system by inducing  
 CC an effector response (cytotoxicity, antibody synthesis, phagocytosis)  
 CC of immune system cells, typically for treating, or immunising against;  
 CC cancer, acquired immune deficiency syndrome, lupus erythematosus,  
 CC multiple sclerosis, toxic shock and snake bite, but also for selective  
 CC destruction of autoreactive cells, diagnostically to assay T cell  
 CC receptors and to raise specific antibodies (useful for diagnosis,  
 CC therapy, studying MHC-associated cellular processes and for affinity  
 CC purification). AAY57558 and AAY68186 to AAY68316 are sequences used in  
 CC the exemplification of the present invention.  
 XX  
 SQ Sequence 253 AA;

Query Match 0.5%; Score 7; DB 21; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAGAL 1115  
 Db 80 feaqgal 86  
 |||||

RESULT 122  
 AAY52931  
 ID AAY52931 standard; Peptide; 253 AA.  
 XX AC AAY52931;  
 XX  
 DT 14-FEB-2000 (first entry)  
 XX  
 DE Class II alpha-chain DR-alpha peptide SEQ ID NO:109.  
 XX  
 KW Major histocompatibility complex; MHC class I; MHC class II; antigen;  
 KW immune response; diagnosis; antibody; immunisation; autoimmune disease;  
 KW acquired immune deficiency syndrome; AIDS; cytostatic; dermatological;  
 KW anti-inflammatory; neuroprotective; immunosuppressive; antithyroid;  
 KW vaccine; lupus erythematosus; multiple sclerosis; thyroiditis;  
 KW toxic shock; tumour; snakebite.  
 XX  
 OS Mammalia.  
 XX  
 PN US5976551-A.  
 XX  
 PD 02-NOV-1999.  
 XX  
 PF 07-JUN-1995; 95US-0484905.  
 XX  
 PR 05-DEC-1991; 91US-0801818.  
 PR 15-NOV-1991; 91US-0792473.  
 XX  
 PA (INSP ) INST PASTEUR.  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX  
 PI Kourilsky P, Mottez E, Abastado J;  
 XX  
 DR WPI; 2000-037081/03.  
 XX  
 PT Composition containing an antigen and altered major histocompatibility  
 PT Class II determinant, used to immunize against autoimmune diseases,  
 PT e.g. acquired immune deficiency syndrome  
 XX  
 PS Disclosure; Column 153-156; 96pp; English.  
 XX  
 CC The present invention describes a composition capable of eliciting  
 CC anti-major histocompatibility (MHC) antibodies. The composition

CC comprises an antigen associated with an altered MHC Class II determinant  
 CC (I) comprising alpha1, alpha2, beta1 and beta2 polypeptide domains.  
 CC encoded by a mammalian MHC Class II locus covalently linked to form a  
 CC polypeptide (I) containing beta2, alpha2, alpha1 and beta1 domains in  
 CC sequence. The resulting Antigen-MHC complex is recognizable by the T cell  
 CC receptor. The compositions are used for immunisation against, or  
 CC treatment of, a wide range of autoimmune diseases, e.g. acquired immune  
 CC deficiency syndrome (AIDS), lupus erythematosus, multiple sclerosis,  
 CC thyroiditis, toxic shock, tumour and snakebite, depending on the nature  
 CC of antigen. (I) is also used to analyse functional interactions between  
 CC the various domains and for targeting lymphocyte receptors. Antibodies  
 CC against (I) are produced by usual methods of immunisation or cell fusion,  
 CC and may be humanised by standard methods. These antibodies are useful for  
 CC diagnosis (detection or purification of MHC gene products), therapy  
 CC (neutralising MHC on cell surfaces) and in the study of MHC and cellular  
 CC processes. AA33240 to AA33242 and AA52840 to AA52970 represent  
 CC sequences used in the exemplification of the present invention.

XX  
 SQ Sequence 253 AA;

Query Match 0.5%; Score 7; DB 21; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115

Db 80 feaqal 86  
 |||||

RESULT 123

AA58692  
 ID AAB58692 standard; protein; 253 AA.

XX  
 AC AAB58692;

DT 13-MAR-2001 (first entry)

DE Class II alpha chain protein #1.

XX Major histocompatibility complex; MHC class I; immune; snake bite;  
 KW T cell mediated autoimmune disease; AIDS; lupus erythematosus;  
 KW toxic shock.

XX Unidentified.

OS US6153408-A.

XX 28-NOV-2000.

XX 09-JAN-1995; 95US-0370476.

XX 15-NOV-1991; 91US-0792473.

XX 07-SEP-1993; 93US-0117575.

XX 05-DEC-1991; 91US-0801818.

XX 07-JUN-1993; 93US-0072787.

XX (INSP ) INST PASTEUR.

PA (INRM ) INST NAT SANTE & RECH MEDICAL.

XX Abastado J, Kourilsky P, Casrouge A, Ojcius D, Lone Y, Mottez E;

XX WPI; 2001-060089/07.

XX New altered major histocompatibility complex (MHC) class I determinant  
 PT useful for eliciting an immune response and/or for immunizing against  
 PT or treating diseases, for example, multiple sclerosis, AIDS, toxic  
 PT shock or snake bite

XX Disclosure; Column 35-36; 105pp; English.

XX The present invention relates to a major histocompatibility complex  
 CC (MHC) class I determinant, which has alpha\_1 alpha\_2 alpha\_3 and

CC beta2-microglobulin polypeptide domains encoded by a mammalian MHC  
 CC class I locus. The MHC class I determinants are useful for activating  
 CC the immune system and presenting antigens to the immune system to  
 CC elicit an antigenic response. The MHC class I determinants are also  
 CC useful for treating diseases, e.g. T cell mediated autoimmune disease,  
 CC AIDS, lupus erythematosus, toxic shock or snake bite. The altered MHC  
 CC class I determinants and compositions containing antigens bound to  
 CC the determinants are useful in diagnostic applications, e.g. altered  
 CC determinants may be used to target lymphocyte receptors and the  
 CC resulting bound determinant can be assayed.

XX Sequence 253 AA;

Query Match 0.5%; Score 7; DB 22; Length 253;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115

Db 80 feaqal 86  
 |||||

RESULT 124

AA58282  
 ID AAY68282 standard; Peptide; 256 AA.

XX  
 AC AAY68282;

DT 13-APR-2000 (first entry)

DE Class II alpha chain protein I-3-alpha SEQ ID NO:114.

XX MHC class I; major histocompatibility complex; microglobulin; antigen;  
 KW immune response; immunisation; AIDS; multiple sclerosis; toxic shock;  
 KW cancer; lupus erythematosus; snake bite; cytostatic; antiviral;  
 KW immunomodulatory; dermatological; immunosuppressive; antiinflammatory;  
 KW neuroprotective.

XX Unidentified.

XX US6011146-A.

XX 04-JAN-2000.

XX 07-JUN-1995; 95US-0481985.

XX 15-NOV-1991; 91US-0792473.

XX 05-DEC-1991; 91US-0801818.

XX (INSP ) INST PASTEUR.

PA (INRM ) INST NAT SANTE & RECH MEDICALE.

XX Kourilsky P, Mottez E, Abastado J;

XX WPI; 2000-125951/11.

XX New recombinant DNA encoding covalently linked form of major  
 PT histocompatibility complex Class I determinant, used for immune system  
 PT stimulation, e.g. for treating cancer

XX Disclosure; Column 137-140; 88pp; English.

XX The present invention describes a recombinant DNA molecule (I)  
 CC containing a sequence (Ia) that encodes an altered MHC (major  
 CC histocompatibility complex ) Class I determinant (II) comprises a  
 CC polypeptide with alpha1, alpha2, alpha3 and beta2-microglobulin  
 CC domains, in which alpha3 and beta2 are covalently linked, thorough C-  
 CC and N-termini respectively, via a nucleotide spacer sequence encoding a  
 CC polypeptide. (II) includes an antigen-binding site and when (II) and  
 CC the antigen are associated they are recognized by a mammalian T cell  
 CC receptor (TCR). (I) are used to produce (II) which are used to study  
 CC functional interactions between the various MHC domains. They can also

CC be used to modulate (in vivo or in vitro) the immune system by inducing  
 CC an effector response (cytotoxicity, antibody synthesis, phagocytosis)  
 CC of immune system cells, typically for treating, or immunising against;  
 CC cancer, acquired immune deficiency syndrome, lupus erythematosus,  
 CC multiple sclerosis, toxic shock and snake bite, but also for selective  
 CC destruction of autoreactive cells, diagnostically to assay T cell  
 CC receptors and to raise specific antibodies (useful for diagnosis,  
 CC therapy, studying MHC-associated cellular processes and for affinity  
 CC purification). AAY57358 and AAY68186 to AAY68316 are sequences used in  
 CC the exemplification of the present invention.

XX  
 SQ Sequence 256 AA;

Query Match 0.5%; Score 7; DB 21; Length 256;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115

Db 80 feaqal 86  
 |||||

RESULT 125

AAY52936

ID AAY52936 standard; Peptide; 256 AA.

XX  
 AC AAY52936;

XX  
 DT 14-FEB-2000 (first entry)

DE Class II alpha-chain I-E-alpha peptide SEQ ID NO:114.

XX Major histocompatibility complex; MHC class I; MHC class II; antigen;  
 KW immune response; diagnosis; antibody; immunisation; autoimmune disease;  
 KW acquired immune deficiency syndrome; AIDS; cytostatic; dermatological;  
 KW anti-inflammatory; neuroprotective; immunosuppressive; antithyroid;  
 KW vaccine; lupus erythematosus; multiple sclerosis; thyroiditis;  
 KW toxic shock; tumour; snakebite.

XX Mammalia.

XX US5976551-A.

PD 02-NOV-1999.

XX PF 07-JUN-1995; 95US-0484905.

XX PR 05-DEC-1991; 91US-0801818.

XX PR 15-NOV-1991; 91US-0792473.

XX PA (INSP ) INST PASTEUR.

PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

XX Kourilsky P, Mottez E, Abastado J;

XX WPI; 2000-037081/03.

XX Composition containing an antigen and altered major histocompatibility  
 PT Class II determinant, used to immunize against autoimmune diseases,  
 PT e.g. acquired immune deficiency syndrome -

PS Disclosure; Column 161-164; 96pp; English.

CC The present invention describes a composition capable of eliciting  
 CC anti-major histocompatibility (MHC) antibodies. The composition  
 CC comprises an antigen associated with an altered MHC Class II determinant  
 CC (I) comprising alpha1, alpha2, beta1 and beta2 polypeptide domains  
 CC encoded by a mammalian MHC Class II locus covalently linked to form a  
 CC polypeptide (I) containing beta2, alpha2, alpha1 and beta1 domains in  
 CC sequence. The resulting Antigen-MHC complex is recognizable by the T cell  
 CC receptor. The compositions are used for immunisation against, or  
 CC treatment of, a wide range of autoimmune diseases, e.g. acquired immune

CC deficiency syndrome (AIDS), lupus erythematosus, multiple sclerosis,  
 CC thyroiditis, toxic shock, tumour and snakebite, depending on the nature  
 CC of antigen. (I) is also used to analyse functional interactions between  
 CC the various domains and for targeting lymphocyte receptors. Antibodies  
 CC against (I) are produced by usual methods of immunisation or cell fusion,  
 CC and may be humanised by standard methods. These antibodies are useful for  
 CC diagnosis (detection or purification of MHC gene products), therapy  
 CC (neutralising MHC on cell surfaces) and in the study of MHC and cellular  
 CC processes. AAY33240 to AAY33242 and AAY52840 to AAY52970 represent  
 CC sequences used in the exemplification of the present invention.

XX  
 SQ Sequence 256 AA;

Query Match 0.5%; Score 7; DB 21; Length 256;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115

Db 80 feaqal 86  
 |||||

RESULT 126

AAB58697

ID AAB58697 standard; protein; 256 AA.

XX  
 AC AAB58697;

XX DT 13-MAR-2001 (first entry)

DE Class II alpha chain protein #6.

XX Major histocompatibility complex; MHC class I; immune; snake bite;  
 KW T cell mediated autoimmune disease; AIDS; lupus erythematosus;  
 KW toxic shock.

XX OS Unidentified.

XX PN US6153408-A.

XX PD 28-NOV-2000.

XX PF 09-JAN-1995; 95US-0370476.

XX PR 15-NOV-1991; 91US-0792473.

XX PR 07-SEP-1993; 93US-0117575.

XX PR 05-DEC-1991; 91US-0801818.

XX PR 07-JUN-1993; 93US-0072787.

XX PA (INSP ) INST PASTEUR.

PA (INRM ) INST NAT SANTE & RECH MEDICAL.

XX Abastado J, Kourilsky P, Casrouge A, Ojcius D, Lone Y, Mottez E;

XX WPI; 2001-060089/07.

XX New altered major histocompatibility complex (MHC) class I determinant  
 PT useful for eliciting an immune response and/or for immunizing against  
 PT or treating diseases, for example, multiple sclerosis, AIDS, toxic  
 PT shock or snake bite -

PS Disclosure; Column 35-36; 105pp; English.

CC The present invention relates to a major histocompatibility complex  
 CC (MHC) class I determinant, which has alpha\_1 alpha\_2 alpha\_3 and  
 CC beta2-microglobulin polypeptide domains encoded by a mammalian MHC  
 CC class I locus. The MHC class I determinants are useful for activating  
 CC the immune system and presenting antigens to the immune system to  
 CC elicit an antigenic response. The MHC class I determinants are also  
 CC useful for treating diseases, e.g. T cell mediated autoimmune disease,  
 CC AIDS, lupus erythematosus, toxic shock or snake bite. The altered MHC  
 CC class I determinants and compositions containing antigens bound to

CC the determinants are useful in diagnostic applications, e.g. altered  
 CC determinants may be used to target lymphocyte receptors and the  
 CC resulting bound determinant can be assayed.

XX Sequence 256 AA;

Query Match 0.5%; Score 7; DB 22; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 |||||  
 Db 80 feaqal 86

RESULT 127

AA211120  
 ID AAY211120 standard; Protein; 272 AA.

XX AC AAY211120;

DT 22-JUL-1999 (first entry)

XX Human bcl2 proto-oncogene wild type protein fragment 17.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;  
 KW frameshift mutation; age-related disease; neurodegenerative disorder;  
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;  
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;  
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;  
 KW ubiquitin B; apolipoprotein E; MAR2; neurofilament-L; neurofilament-M;  
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;  
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;  
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;  
 KW high mobility group protein-C; neuroendocrine specific protein A.

XX Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

PF 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI; 1998-609901/51.

XX N-PSDB; AAX75766.

XX Diagnosing disease by detecting frameshift mutations in RNA or  
 PT corresponding protein mutations - used to diagnose cancer and  
 PT neurological diseases, particularly Alzheimer's disease, and also  
 PT for treatment and prevention with specific ribozymes or wild-type  
 RNA

PS Disclosure; Figure 15; 259pp; English.

XX This invention describes a novel method for the diagnosis of a disease  
 CC caused by, or associated with, an RNA molecule that has a frameshift  
 CC mutation. The method is used to diagnose age-related diseases, especially  
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
 CC and many others listed) or susceptibility to these disorders. The method  
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
 CC at an early stage. It is based on the observation that disease may be

CC caused by mutations in RNA rather than DNA. The invention describes the  
 CC used of neuronal system RNA molecules, specifically proteins including  
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
 CC proteins tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,  
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic  
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group  
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 272 AA;

Query Match 0.5%; Score 7; DB 19; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22

Db 255 slalvga 261  
 |||||

RESULT 128

AA58160

ID AAB58160 standard; Protein; 281 AA.

XX AC AAB58160;

DT 14-MAR-2001 (first entry)

DE Lung cancer associated polypeptide sequence SEQ ID 498.

XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardiointestine; immunomodulatory; muscular active; vulnary;  
 KW gastrointestinal; nephrotropic; antinefactive; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease.

XX Homo sapiens.

XX WO200055180-A2.

XX 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05918.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI; 2000-587514/55.

XX N-PSDB; AAF18036.

XX Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer

PS Claim 11; Page 986-987; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytostatic; cardioactive;  
 CC immunomodulatory; muscular active general; vulnary; gastrointestinal  
 CC general; nephrotropic; antinefactive; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat



CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.

XX Sequence 281 AA;

Query Match 0.5%; Score 7; DB 21; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1109 FEAQAL 1115

DB 106 feaqal 112  
 |||||

RESULT 129

AAFP70392  
 ID AAF70392 standard; Protein; 283 AA.

XX AAF70392;

XX 26-APR-1991 (first entry)

XX Alpha-subunit of Sarcophaga lectin.

XX Antitumour; cancer.

XX Sarcophaga peregrina.

XX Key Location/Qualifiers

XX Peptide 1..23

XX /label= Presumed signal peptide

XX JP62029981-A.

XX 07-FEB-1987.

XX 23-JUL-1985; 85JP-0161058.

XX 23-JUL-1985; 85JP-0161058.

XX 01-JAN-1985; 85JP-0233516.

XX (WAKU-) WAKUNAGA SEIYAKU KK.

XX (SAGA-) SAGAMI HAM K.

XX (TEIJ ) TEIJIN KK.

XX (YOSH ) YOSHITOMO PHARM IND KK.

XX WPI; 1987-076448/11.

XX N-PSDB; AAN70615.

XX Single-stranded DNA contg. nucleotide(s) - and for encoding  
 PT aminoacid of sarcophaga lectin alpha subunits from injured  
 PT peregrina larva

XX Disclosure; Fig 3; 9pp; Japanese.

XX Sarcophaga lectin has antitumour activity, therefore alpha-sununit  
 CC gene product is assumed to have similar activity.  
 XX Sequence 283 AA;

Query Match 0.5%; Score 7; DB 8; Length 283;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 801 NNPDNYK 807

|||||

DB 121 nnpdnyk 127

RESULT 130

AAFP70427

ID AAF70427 standard; protein; 283 AA.

XX AAF70427;

XX 05-APR-1991 (first entry)

XX Sequence encoding alpha-subunit of Sarcophaga lectin (SL) of

DE Sarcophaga peregrina.

XX Lectin-like protein; antibacterial; antitumour; ss.

XX Sarcophaga peregrina.

XX Key Location/Qualifiers

XX Protein 24..283

XX /note="claimed"

XX JP621111999-A.

XX 22-MAY-1987.

XX 23-JUL-1985; 85JP-0233516.

XX 01-JAN-1985; 85JP-0233516.

XX 23-JUL-1985; 85JP-0161058.

XX (TEIJ ) TEIJIN KK.

XX WPI; 1987-181911/26.

XX P-PSDB; AAF70427.

XX Physiologically active protein - comprises amino acid sequence  
 PT comprising alpha sub-unit of Sarcophaga lectin, used as  
 PT anti-bacterial agent.  
 XX Disclosure; Fig 3; 9pp; Japanese.

XX SL is a protein having lectin-like activity which is yielded in body  
 CC humor of larva of Sarcophaga peregrina on injuring the surface of  
 CC the body. It is useful as antibacterial agent and anti-tumour drug.

XX Sequence 283 AA;

Query Match 0.5%; Score 7; DB 8; Length 283;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 801 NNPDNYK 807

|||||

DB 121 nnpdnyk 127

RESULT 131

AAB32503

ID AAB32503 standard; Protein; 290 AA.

XX AAB32503;

XX 19-JAN-2001 (first entry)

XX S. lavendulae Mit I encoded protein sequence.

XX Mitomycin; biosynthesis; mitosane ring system; antibiotic; anti-cancer;  
 XX anti-inflammatory; immune-enhancer; immunosuppressant; asthma;  
 XX chronic obstructive pulmonary disease; respiratory inflammation;  
 XX fungicide; pesticide.

## OS Streptomyces lavendulae.

XX WO200053737-A2.  
 XX PD 14-SEP-2000.  
 XX 10-MAR-2000; 2000WO-US06394.  
 XX PF 12-MAR-1999; 99US-0266965.  
 XX PA (MINU ) UNIV MINNESOTA.  
 XX PA (SHER/) SHERMAN D H.  
 XX PA (MAOY/) MAO Y.  
 XX PA (VARO/) VAROGLU M.  
 XX PA (HEMM/) HE M.  
 XX PA (SHEL/) SHELTON P C.  
 XX PI Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;  
 XX DR WPI; 2000-601980/57.  
 XX DR N-PSDB; AAC55806.  
 XX PT Novel nucleic acid molecule comprising mitomycin biosynthetic gene  
 XX PT cluster useful for cloning mitomycin biosynthetic genes for elucidating  
 XX PT the molecular basis of mitosane ring system biosynthesis  
 XX PS Disclosure; Page 347-348; 399pp; English.

XX This invention relates to isolated and purified nucleic acid molecules  
 CC from the mitomycin biosynthetic gene cluster. Mitomycins are a group of  
 CC natural products that contain a variety of functional groups, including  
 CC amino benzoquinone and axiridine ring systems. The S. lavendulae  
 CC mitomycin biosynthetic gene cluster comprises 47 mitomycin genes  
 CC spanning 55kb of DNA. The invention includes an expression cassette  
 CC comprising a mitomycin biosynthetic gene operably linked to a promoter,  
 CC and host cells transformed with the cassette. The nucleotide, and protein  
 CC sequences and the transformed host cells of the invention result in  
 CC antitumor, antiinflammatory, cytostatic, immunomodulatory, and  
 CC antibiotic activities. The nucleotide sequences are used to elucidate the  
 CC molecular basis for the biosynthesis of novel natural products, as well  
 CC as to engineer the biosynthesis of novel natural products, e.g.  
 CC antibiotics, anti-inflammatory agents, anti-cancer agents,  
 CC immune-enhancers, immunosuppressants, agents to treat asthma, chronic  
 CC obstructive pulmonary disease as well as other disease involving  
 CC respiratory inflammation, or cholesterol-lowering agents or as crop  
 CC protection agents (e.g. fungicides or insecticides) as well as  
 CC biopolymers, e.g., in packaging or biomedical applications, or to engineer  
 CC PHA monomer synthases. Sequences AAC55782-C55881, AAC55815-C55849 and  
 CC AAB32485-B32542 represent mitomycin biosynthetic gene cluster DNA  
 CC sequences and encoded proteins. Sequences AAC55812-C55814,  
 CC AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the  
 CC cloning of the mitomycin biosynthetic genes.

XX Sequence 290 AA;

Query Match 0.5%; Score 7; DB 21; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 207 VGSAGR 213  
 |||||

Db 43 vgsagr 49

## RESULT 132

XX AAY86120  
 ID AAY86120 standard; Protein; 293 AA.

XX AC AAY86120;

XX 10-APR-2000 (first entry)

XX

## DE S. pneumoniae derived protein #329.

XX Treatment; prevention; disease; diagnosis; gene therapy; screening;  
 KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.  
 XX Streptococcus pneumoniae.  
 XX WO9806734-A1.  
 XX PD 19-FEB-1998.  
 XX PF 15-AUG-1997; 97WO-US14436.  
 XX PR 16-AUG-1996; 96US-0024022.  
 XX PA (SMIK ) SMITHLINE BEECHAM CORP.  
 XX PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;  
 XX PI Stodola RK;  
 XX DR WPI; 1998-159452/14.  
 XX DR N-PSDB; AAZ96436.  
 XX PT Streptococcus pneumoniae proteins and related DNA - useful for  
 XX PT screening compounds for antibacterial activity  
 XX PS Claim 5; Page 589-590; 640pp; English.

XX This invention describes novel isolated Streptococcus pneumoniae  
 CC polynucleotides (see AAZ96173-296494) and their encoded proteins (see  
 CC AAY85792-186182). The DNA, vectors and host cells described in the  
 CC method of the invention are useful for the recombinant expression of the  
 CC polypeptides. The polypeptides are useful for treatment or prevention of  
 CC disease, or diagnosis of disease related to expression or activity of  
 CC such a polypeptide. They can also be used to screen for compounds which  
 CC interact with and inhibit or activate such a polypeptide. The  
 CC polypeptides (or DNA encoding them, via gene therapy) are also useful  
 CC for inducing an immunological response in a mammal. The antagonists are  
 CC useful to inhibit such bacterial polypeptides. The polypeptides are  
 CC particularly useful to identify antimicrobial compounds and antibiotics.  
 CC They are also useful to determine their role in pathogenesis of  
 CC infection, dysfunction and disease.

XX Sequence 293 AA;

Query Match 0.5%; Score 7; DB 19; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1261 LKLAKEV 1267  
 |||||

Db 116 lklakev 122

## RESULT 133

XX AAY10990  
 ID AAY10990 standard; Protein; 294 AA.

XX AC AAY10990;

XX 08-JUN-1999 (first entry)

DE H. pylori ORF 06ep11202\_4569693\_c2\_28 cellular protein.

XX Vaccine; probe; diagnostic; ORF; cell envelope protein;  
 KW secreted protein; cellular protein.

XX Helicobacter pylori.

XX WO9818323-A1.

XX 07-MAY-1998.

XX PF 28-OCT-1997; 97WO-US19575.  
 XX PR 14-JUL-1997; 97US-0891928.  
 XX PR 28-OCT-1996; 96US-0739150.  
 XX PR 06-DEC-1996; 96US-0759739.  
 XX PA (ASTR ) ASTRA AB.  
 XX PI Alm RA, Smith D;  
 XX WPI; 1998-271811/24.  
 XX DR N-PSDB; AAX30457.  
 XX Helicobacter pylori nucleic acids and proteins - used to develop  
 PT products for the detection, prevention and treatment of H. pylori  
 PT infections  
 XX Claims 27, 31; Page 195; 279pp; English.  
 XX Recombinant or substantially pure preparations of H. pylori polypeptides  
 CC are disclosed, together with the nucleic acids encoding them. In all,  
 CC 73 ORFs are shown. The proteins are variously cell envelope proteins,  
 CC secreted proteins or other cellular proteins. Vaccines containing the  
 CC nucleic acids or proteins are claimed, as are probes containing at least  
 CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful  
 CC for treating or reducing the risk of H. pylori infections, and the  
 CC probes can be used diagnostically for detecting the presence of  
 CC Helicobacter in a sample. The products are also of use in screening  
 CC for compounds having the ability to interfere with the H. pylori life  
 CC cycle or to inhibit H. pylori infection.  
 XX SQ Sequence 294 AA;

Query Match 0.5%; Score 7; DB 19; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 954 NNIASLE 960  
 Db 217 nniase 223  
 |||||

RESULT 134  
 AAW20738  
 ID AAW20738 standard; protein: 297 AA.  
 AC AAW20738;

DT 16-JUL-1997 (first entry)

DE H. pylori cytoplasmic protein, 06cp20302orf8.

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.

OS Helicobacter pylori.

PN W09640893-A1.

PD 19-DEC-1996.

PF 06-JUN-1996; 96WO-US09122.

PR 01-APR-1996; 96US-0630405.

PR 07-JUN-1995; 95US-0487032.

XX (ASTR ) ASTRA AB.

XX Berglindh OT, Smith D, Mellgaard BL;

XX

DR WPI; 1997-052306/05.  
 DR N-PSDB; AAT67991.

XX Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter  
 XX Claim 61; Page 1154; 1481pp; English.

PS The present sequence is a Helicobacter pylori cytoplasmic protein.  
 XX The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.

CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.

XX SQ Sequence 297 AA;

Query Match 0.5%; Score 7; DB 18; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 954 NNIASLE 960  
 Db 220 nniase 226  
 |||||

RESULT 135

AAAR82538

ID AAR82538 standard; protein: 298 AA.

XX AAR82538;

DT 17-APR-1996 (first entry)

XX Hybrid IA beta chain.

XX Polymerase chain reaction; PCR; primer; amplify;

KW major histocompatibility complex; MHC; T-cell receptor; TCR;

KW autoimmune disease; immunodeficiency disease; immune response;

KW immunoproliferation disease; graft-host rejection; therapy; B cell;  
 KW M12.C3; pM12-1AB-Ea.

XX Synthetic.

XX Key Location/Qualifiers

FT Region 1..27

FT /note= "leader region"

FT Misc-difference 28..31

FT /note= "IA beta chain beta 1 domain fragment"

FT Peptide 32..48

FT /note= "IE alpha chain peptide fragment"

FT Region 49..63

FT /note= "linker region"

FT Domain 64..156

FT /note= "IA beta chain beta 1 domain"

FT Domain 157..298

FT /note= "IA beta chain beta 2 domain"

FT Domain 260..280

FT /note= "Transmembrane domain"

FT Domain 281..298

FT /note= "cytoplasmic domain"

XX W09523814-A1.

PN

XX PD 08-SEP-1995.  
 XX PF 03-MAR-1995; 95WO-US02689.  
 XX PR 04-MAR-1994; 94US-0207481.  
 XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.  
 XX Kappler JW, Marrack P;  
 XX WPI; 1995-320543/41.  
 XX N-PSDB; AAT04269.  
 XX Peptide-MHC complex comprising antigenic peptide, linker and MHC  
 XX segment - useful as reagents for the treatment of diseases including  
 XX auto-immune diseases, immuno-stimulatory diseases or graft-host  
 XX rejection  
 XX Example 2; Page 65; 94pp; English.  
 XX This sequence represents a hybrid IA beta chain. This sequence  
 XX contains a fragment of the IE alpha chain (residues 56-73), as well as a  
 XX linker and cleavage site. The DNA encoding this sequence was transfected  
 XX into a B cell line (MI2.C3) using plasmid pMI2-TAB-Ea. It was found that  
 XX this sequence was expressed in these cells. Complexes such as this may  
 XX be used to regulate an immune response. The complexes are capable of  
 XX being recognised by a TCR alone or in combination with additional MHC  
 XX proteins. These complexes are useful for therapeutic purposes and  
 XX experimental purposes. They can also be used as reagents for the  
 XX treatment of diseases including autoimmune diseases, immunodeficiency  
 XX diseases, immunoproliferation diseases, and graft-host rejection.  
 XX Sequence 298 AA;  
 SQ

Query Match 0.5%; Score 7; DB 16; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 FEAQAL 1115  
 |||||  
 Db 34 feaqal 40

RESULT 136  
 AAW55450  
 ID AAW55450 standard; Protein; 299 AA.  
 AC AAW55450;  
 XX  
 DT 24-JUN-1998 (first entry)  
 XX  
 DE H. pylori ORF 02ael1612\_22477267\_f2\_27 cell envelope OMP.  
 XX  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9737044-A1.  
 XX  
 PD 09-OCT-1997.  
 XX  
 XX 27-MAR-1997; 97WO-US05223.  
 XX  
 PR 06-DEC-1996; 96US-0761318.  
 PR 29-MAR-1996; 96US-0625811.  
 PR 02-APR-1996; 96US-0758731.  
 PR 25-OCT-1996; 96US-0736905.  
 PR 28-OCT-1996; 96US-0738859.  
 XX

PA (ASTR ) ASTRA AB.  
 XX Alm RA, Smith D;  
 XX WPI; 1997-503122/46.  
 DR N-PSDB; AAV24859.  
 XX  
 PT Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX  
 PS Claims 14,80; Page 657; 1145pp; English.  
 XX  
 CC This sequence is a H. pylori cell envelope outer membrane protein  
 CC (OMP) having no terminal Phe residue.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.  
 XX  
 SQ Sequence 299 AA;  
 Query Match 0.5%; Score 7; DB 18; Length 299;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGAL 23  
 |||||  
 Db 8 lalvgal 14

RESULT 137  
 AAW98322  
 ID AAW98322 standard; Protein; 299 AA.  
 AC AAW98322;  
 XX  
 DT 31-MAR-1999 (first entry)  
 XX  
 DE H. pylori GHPO 1360 protein.  
 XX  
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9843478-A1.  
 XX  
 PD 08-OCT-1998.  
 XX  
 PF 01-APR-1998; 98WO-US06371.  
 XX  
 PR 29-JUL-1997; 97US-0902615.  
 PR 01-APR-1997; 97US-0833457.  
 PR 24-JUN-1997; 97US-0881227.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.

PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

XX  
PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
XX WPI; 1998-542293/46.

XX N-PSDB; AAX14041.

XX  
PT New isolated Helicobacter polynucleotides - used to develop products  
PT for the diagnosis, prevention and treatment of Helicobacter  
PT infections and gastrointestinal diseases

XX  
PS Claim 8; Page 431-432; 2054pp; English.

XX  
CC This sequence represents a Helicobacter pylori GHPO protein of the  
CC invention. The polypeptides can be used for preventing or treating  
CC Helicobacter infections, and gastroduodenal diseases associated with  
CC these infections, including acute, chronic, and atrophic gastritis, and  
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
CC used for the production of antibodies. The products can also be used for  
CC detection and diagnosis.

XX  
SQ Sequence 299 AA;

Query Match 0.5%; Score 7; DB 19; Length 299;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGAL 23  
Db 8 laivgal 14

RESULT 138

AAW73034  
ID AAW73034 standard; Protein; 299 AA.

XX  
AC AAW73034;

XX  
DT 02-FEB-1999 (first entry)

XX  
DE Helicobacter pylori 32 kDa polypeptide GHPO 1360.

XX  
KW GHPO 1360; infection; gastritis; ulcer; vaccine; diagnosis;  
KW therapy.

XX  
OS Helicobacter pylori.

XX  
FH Key Location/Qualifiers  
FT -Peptide 1..20  
FT Protein /label= Sig\_peptide  
FT 20..299  
FT /label= Mat\_protein

XX  
PN WO9843479-A1.

XX  
PD 08-OCT-1998.

XX  
PF 31-MAR-1998; 98WO-US06421.

XX  
PR 01-APR-1997; 97US-0834666.

XX  
PR 01-APR-1997; 97US-0831310.

XX  
(HUMA-) HUMAN GENOME SCI INC.

PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS ET VACCINS.

XX  
PI Al-Garawi A, Kleanthous H, Lissolo L, Miller C, Tomb J;

XX  
DR WPI; 1998-568251/48.

XX  
DR N-PSDB; AAW07963.

XX  
PT New isolated Helicobacter polynucleotides - used to develop products  
PT for the diagnosis, prevention and treatment of Helicobacter

PT infections and gastroduodenal diseases

XX  
PS Claim 9; Page 148-149; 184pp; English.

XX  
CC This is the amino acid sequence of a 32 kDa Helicobacter pylori  
CC polypeptide designated GHPO 1360. It was deduced from an isolated  
CC genomic DNA sequence (see AAW07963). The invention provides a family  
CC of 76 kDa Helicobacter polypeptides (see AAW73023-32), as well as  
CC GHPO 1360 and a 50 kDa polypeptide (see AAW73035), and also  
CC polynucleotides (see AAW72001, AAW07912-21 and AAW07963-64) encoding or  
CC them, expression cassettes, and methods for producing the unprocessed or  
CC mature polypeptides in host cells. The polypeptides can be used in  
CC vaccination methods to prevent or treat Hb infection in a mammal.  
CC Methods and products of the invention allow treatment and  
CC prevention of gastroduodenal diseases associated with Hb  
CC infections, including acute, chronic, and atrophic gastritis, and  
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. Detection  
CC and diagnostic methods are also provided. GHPO 1360 was  
CC demonstrated to be a protective antigen.

XX  
SQ Sequence 299 AA;

Query Match 0.5%; Score 7; DB 19; Length 299;

Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LALVGAL 23  
Db 8 laivgal 14

RESULT 139

AAW89829  
ID AAW89829 standard; Protein; 299 AA.

XX  
AC AAW89829;

XX  
DT 18-FEB-1999 (first entry)

XX  
DE Protein encoded by clone Y104-1.asm.

XX  
KW Antigen; Immunogenic cluster family; vaccine; gastritis; diagnosis;  
KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma.

XX  
OS Helicobacter pylori.

XX  
PN WO9849314-A2.

XX  
PD 05-NOV-1998.

XX  
PF 27-APR-1998; 98WO-US08487.

XX  
PR 14-OCT-1997; 97US-0061958.

XX  
PR 25-APR-1997; 97US-0045107.

XX  
PA (GENE-) GENELABS TECHNOLOGIES INC.

XX  
PI Chow TP, Fry KE, Lim MY, McAttee CP;

XX  
DR WPI; 1999-009433/01.

XX  
DR N-PSDB; AAV90580.

XX  
PT New Helicobacter pylori antigens and related nucleic acid sequences  
PT - useful in serological diagnosis and protective vaccines, providing  
PT long-lasting immune response

XX  
PS Claim 15; Page 120; 402pp; English.

XX  
CC The present sequence represents a Helicobacter pylori antigenic protein  
CC that is characterised by immunoreactivity with H. pylori-positive  
CC antisera. The proteins are highly immunogenic and induce a long-lasting  
CC immune response that persists even after antimicrobial treatment. In

CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are  
 CC highly sensitive and specific. The specification also describes 69  
 CC previously unrecognised immunogenic cluster families. H. pylori antigens  
 CC are used to detect H. pylori-specific antibodies, for diagnosing  
 CC infection or to confirm eradication of infection, and in vaccines to  
 CC protect against H. pylori infection and related diseases (gastritis,  
 CC peptic ulcer, gastric adenocarcinoma/lymphoma).

XX SQ Sequence 299 AA;

Query Match 0.5%; Score 7; DB 20; Length 299;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 LALVGL 23  
 Db 8 lalvgal 14  
 |||||

RESULT 140

AAW89849  
 ID AAW89849 standard; Protein; 299 AA.

XX AC  
 XX AC AAW89849;

XX DT 18-FEB-1999 (first entry)

XX DE Antigen from cluster 03.

XX KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;  
 XX KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma.

XX OS Helicobacter pylori.

XX FH Key Location/Qualifiers  
 FT Misc-difference 224  
 FT /note= "not specified"  
 FT Misc-difference 232  
 FT /note= "not specified"

XX PN W09849314-A2.

XX PD 05-NOV-1998.

XX PF 27-APR-1998; 98WO-US08487.

XX PR 14-OCT-1997; 97US-0061958.

XX PR 25-APR-1997; 97US-0045107.

XX PA (GENE-) GENELABS TECHNOLOGIES INC.

XX PI Chow TP, Fry KE, Lim MY, McAtee CP;

XX DR WPI; 1999-009433/01.

XX PT New Helicobacter pylori antigens and related nucleic acid sequences  
 XX PT - useful in serological diagnosis and protective vaccines, providing  
 XX PT long-lasting immune response

XX PS Claim 1; Page 211; 402pp; English.

XX CC The present sequence represents a Helicobacter pylori antigenic protein  
 XX CC that is characterised by immunoreactivity with H. pylori-positive  
 XX CC antisera. The proteins are highly immunogenic and induce a long-lasting  
 XX CC immune response that persists even after antimicrobial treatment. In  
 XX CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are  
 XX CC highly sensitive and specific. The specification also describes 69  
 XX CC previously unrecognised immunogenic cluster families. H. pylori antigens  
 XX CC are used to detect H. pylori-specific antibodies, for diagnosing  
 XX CC infection or to confirm eradication of infection, and in vaccines to  
 XX CC protect against H. pylori infection and related diseases (gastritis,  
 XX CC peptic ulcer, gastric adenocarcinoma/lymphoma).

XX SQ Sequence 299 AA;

Query Match 0.5%; Score 7; DB 20; Length 299;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 LALVGL 23  
 Db 8 lalvgal 14  
 |||||

RESULT 141

AAV00875  
 ID AAV00875 standard; Protein; 305 AA.

XX AC  
 XX AC AAV00875;

XX DT 21-MAY-1999 (first entry)

XX DE Synaptojanin isoform NSYN-1 protein sequence.

XX KW Synaptojanin; human; NSYN-1; neurological disorder; immune disorder;  
 XX KW cancer; therapy.

XX OS Homo sapiens.

XX PN W09906572-A1.

XX PD 11-FEB-1999.

XX PF 28-JUL-1998; 98WO-US15782.

XX PR 31-JUL-1997; 97US-0904234.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Lal P, Tang TY;

XX DR WPI; 1999-153801/13.

XX DR N-PSDB; AAX27072.

XX PT A new synaptojanin isoform - useful to treat neurological disorders,  
 XX PT immune disorders, and cancer

XX PS Claim 1; Fig 1; 41pp; English.

XX CC This sequence is the human synaptojanin isoform NSYN-1 of the  
 XX CC invention. Synaptojanin is used to treat neurological disorders, and  
 XX CC antagonist of it can be used to treat immune disorders or cancers. The  
 XX CC disorders include neurological disorders such as Alzheimer's disease,  
 XX CC amnesia, amyotrophic lateral sclerosis, bipolar disorder, catatonia,  
 XX CC cerebral neoplasms, dementia, depression, Down's syndrome, tardive  
 XX CC dyskinesia, dystonias, epilepsy, Huntington's disease, multiple  
 XX CC sclerosis, neurofibromatosis, Parkinson's disease, schizophrenia, and  
 XX CC Tourette's disorder; cancers of the adrenal gland, bladder, bone,  
 XX CC brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract,  
 XX CC heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis,  
 XX CC prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and  
 XX CC uterus; and immune disorders such as AIDS, Addison's disease, adult  
 XX CC respiratory distress syndrome, allergies, anaemia, asthma,  
 XX CC atherosclerosis, bronchitis, cholecystitis, Crohn's disease, ulcerative  
 XX CC colitis, atopic dermatitis, dermatomyositis, diabetes mellitus,  
 XX CC emphysema, erythema nodosum, atrophic gastritis, glomerulonephritis, lupus  
 XX CC gout, Graves' disease, hyperosinophilia, irritable bowel syndrome, myo-  
 XX CC sarcoma, multiple sclerosis, myosclerosis, myosclerosis, myocardial or  
 XX CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,  
 XX CC polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and  
 XX CC autoimmune thyroiditis.

XX SQ Sequence 305 AA;

Query Match 0.5%; Score 7; DB 20; Length 305;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VKTNGIS 566  
Db 36 vktngis 42

RESULT 142

AAG06505  
ID AAG06505 standard; Protein; 316 AA.

XX AC AAG06505;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 3302.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0128264.

PR 29-MAR-1999; 99US-0128785.

PR 01-APR-1999; 99US-0127462.

PR 08-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0129845.

PR 21-APR-1999; 99US-0130077.

PR 23-APR-1999; 99US-0130449.

PR 28-APR-1999; 99US-0130510.

PR 30-APR-1999; 99US-0130891.

PR 04-MAY-1999; 99US-0132048.

PR 05-MAY-1999; 99US-0132407.

PR 06-MAY-1999; 99US-0132484.

PR 07-MAY-1999; 99US-0132485.

PR 11-MAY-1999; 99US-0132486.

PR 14-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-013256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 24-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 27-MAY-1999; 99US-0136021.

PR 28-MAY-1999; 99US-0136392.

PR 01-JUN-1999; 99US-0136782.

PR 03-JUN-1999; 99US-0137222.

PR 04-JUN-1999; 99US-0137528.

PR 07-JUN-1999; 99US-0137502.

PR 08-JUN-1999; 99US-0137724.

PR 10-JUN-1999; 99US-0138094.

99US-0138540.

PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 18-OCT-1999; 99US-0159638.  
 PR 21-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 22-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161820.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 316;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 456 VDGPLRV 462  
 |||||

Db 84 vdgplrv 90  
 RESULT 143  
 AAG40468  
 ID AAG40468 standard; Protein; 317 AA.  
 XX  
 AC AAG40468;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 50215.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 PN EPI033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126284.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 19-MAY-1999; 99US-0134768.  
 PR 20-MAY-1999; 99US-0134941.  
 PR 21-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 25-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 28-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 03-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 07-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.



PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143325.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144634.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.58; Score 7; DB 21; Length 317;  
Best Local Similarity 100.0%; Pred. No. 3.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 NTDDIKA 789  
| | | | |  
Db 11 ntddika 17

RESULT 144  
AAG06504  
ID AAG06504 standard; Protein; 323 AA.  
XX

AC AAG06504;  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3301.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145145.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145191.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 18-OCT-1999; 99US-0159638.  
 PR 21-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 26-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161356.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 456 VDGPLRV 462  
 Db 91 vdgplrv 97

RESULT 145  
 AAB67352  
 ID AAB67352 standard; protein; 330 AA.  
 XX  
 AC AAB67352;  
 XX

DT 24-APR-2001 (first entry)  
 DE Protein encoded by ALAD coding region of pBlueAaid-3.  
 XX Heme biosynthetic pathway; gene therapy; AIP; ALA;  
 KW

KW delta-aminolevulinic acid; deficient porphyria; ADP;  
 KW porphyria cutanea tarda; PCT; hereditary coproporphyria; HCP;  
 KW harderoporphyria; HDP; variegata porphyria; VP;  
 KW congenital erythropoietic porphyria; CEP;  
 KW erythropoietic protoporphyria; EPP;  
 KW hepaterythropoietic porphyria; HEP.  
 XX Homo sapiens.  
 OS  
 XX WO200107065-A2.  
 PN  
 XX 01-FEB-2001.  
 PD  
 XX 27-JUL-2000; 2000WO-DK00425.  
 PF  
 XX 27-JUL-1999; 99DK-0001071.  
 PR 19-APR-2000; 2000DK-0000667.  
 XX (HEME-) HEMEBIOTECH AS.  
 PA  
 XX Gellerfors P, Fogh J;  
 PI  
 XX WPI; 2001-159639/16.  
 DR  
 XX Treatment or prevention of porphyria, by enzyme replacement or gene  
 PT therapy for correction of mutations, particularly in the  
 PT porphobilinogen deaminase gene  
 XX  
 PS Disclosure; Page 207; 207pp; English.  
 XX  
 CC The present invention relates to treatment or prevention of a  
 CC disease caused by deficiency of at least one enzyme of the  
 CC heme biosynthetic pathway by administering at least one catalyst,  
 CC optionally combined with gene therapy of the relevant mutation.  
 CC The invention is useful for treating and/or preventing AIP, ALA  
 CC (delta-aminolevulinic acid) deficient porphyria (ADP), porphyria  
 CC cutanea tarda (PCT), hereditary coproporphyria (HCP), harderoporphyria  
 CC (HDP), variegata porphyria (VP), congenital erythropoietic porphyria  
 CC (CEP), erythropoietic protoporphyria (EPP) and hepaterythropoietic  
 CC porphyria (HEP).  
 XX  
 SQ Sequence 330 AA;

Query Match 0.5%; Score 7; DB 22; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 TATTTLN 954  
 Db 21 tatttln 27

RESULT 146  
 AAW55731  
 ID AAW55731 standard; Protein; 340 AA.  
 XX  
 AC AAW55731;  
 XX  
 DT 13-JUL-1998 (first entry)  
 XX  
 DE H. pylori ORF 14cpl1908\_24218954\_cl\_68 cell envelope OMP.

XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;  
 KW cell envelope outer membrane protein; OMP; tyrosine cluster motif.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO9737044-A1.  
 XX  
 PD 09-OCT-1997.

XX PF 27-MAR-1997; 97WO-US05223.  
 XX PR 06-DEC-1996; 96US-0761318.  
 XX PR 29-MAR-1996; 96US-0625811.  
 XX PR 02-APR-1996; 96US-0758731.  
 XX PR 25-OCT-1996; 96US-0736905.  
 XX PR 28-OCT-1996; 96US-0738859.  
 XX (ASTR ) ASTRA AB.

XX PA Alm RA, Smith D;  
 XX PI WPI; 1997-503122/46.  
 XX DR N-PSDB; AAV25140.  
 XX XX

PT Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s). - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 XX

PS Claims 14,80; Page 1002,1003; 1145pp; English.

XX This sequence is a H. pylori cell envelope outer membrane protein  
 CC having a terminal Phe residue and a C-terminal tyrosine cluster motif.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds.  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.

XX SQ Sequence 340 AA;

Query Match 0.5%; Score 7; DB 18; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 405 TNADGTI 411  
 Db 219 tnadgti 225  
 |||||

RESULT 147  
 AAB58310  
 ID AAB58310 standard; Protein; 340 AA.  
 XX AC AAB58310;  
 XX DT 14-MAR-2001 (first entry)

XX Lung cancer associated polypeptide sequence SEQ ID 648.  
 DE Human; lung cancer associated protein; neuroprotective; cytostatic;  
 XX cardioactive; immunomodulatory; muscular active; vulnerable;  
 XX gastrointestinal; nephrotropic; antineoplastic; gynecological;  
 XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 XX proliferative disorder; wound healing; infectious disease.  
 XX Homo sapiens.

XX PN WO200055180-A2.  
 XX PD 21-SEP-2000.  
 XX PF 08-MAR-2000; 2000WO-US05918.  
 XX PR 12-MAR-1999; 99US-0124270.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX PA (ROSE/) ROSEN C A.  
 XX PI Ruben SM;  
 XX XX

XX WPI; 2000-587514/55.  
 XX DR N-PSDB; AAF18186.  
 XX XX

PT Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 XX

PS Claim 11; Page 1145-1146; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytostatic; cardioactive;  
 CC immunomodulatory; muscular active general; vulnerable; gastrointestinal  
 CC general; nephrotropic; antineoplastic; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.

XX SQ Sequence 340 AA;

Query Match 0.5%; Score 7; DB 21; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 966 LQTLSLS 972  
 Db 60 lqtlsls 66  
 |||||

RESULT 148  
 AAB15928  
 ID AAB15928 standard; Protein; 344 AA.  
 XX AC AAB15928;  
 XX DT 05-OCT-2000 (first entry)

XX E. coli proliferation associated protein sequence SEQ ID NO:285.  
 XX Escherichia coli; E. coli; proliferation; inhibition; screening;  
 XX antimicrobial; bacterial growth; antisense therapy; antibacterial.  
 XX Escherichia coli.  
 XX WO200044906-A2.  
 XX PD 03-AUG-2000.

PF 27-JAN-2000; 2000WO-US02200.  
 XX  
 PR 27-JAN-1999; 99US-0117405.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Zyskind J, Ohlsen KL, Trawick J, Forsyth RA, Froelich JM, Carr GU;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI; 2000-514822/46.  
 DR N-PSDB; AAA65933.  
 XX  
 XX Novel polynucleotides and polypeptides associated with microorganism  
 PT proliferation, used to identify inhibitors of bacterial growth and  
 PT proliferation, for use in antisense therapy -  
 XX  
 PS Claim 11; Page 212; 316pp; English.  
 XX  
 CC AAA65809 to AAA65989 and AAA66058 to AAA66138 represent nucleotide  
 CC sequences derived from *Escherichia coli* which inhibit *E. coli*  
 CC proliferation. AAA65890 to AAA66055 and AAA65886 to AAA66040 represent  
 CC nucleotide and protein sequences associated with *E. coli* proliferation.  
 CC AAA66056 and AAA66057 represent primers used for sequencing *E. coli*  
 CC proliferation inhibiting nucleotide inserts in an example from the  
 CC present invention. Methods from the present invention can be used to  
 CC identify a proliferation required gene in a microorganism, by contacting  
 CC a microorganism with a proliferation-required gene activity inhibitory  
 CC nucleic acid identified in another organism, and determining if  
 CC inhibition occurs in the second microorganism. The nucleic acid sequences  
 CC identified as being required for bacterial growth and proliferation, can  
 CC be used for antisense therapy for killing bacteria.  
 XX  
 SQ Sequence 344 AA;

Query Match 0.5%; Score 7; DB 21; Length 344;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 861 NALAQNA 867  
 |||||  
 DB 172 nalaqna 178

RESULT 149  
 AAY75482  
 ID AAY75482 standard; Protein; 351 AA.  
 XX  
 AC AAY75482;  
 XX  
 DT 21-MAR-2000. (first entry)  
 XX  
 DE *Neisseria meningitidis* ORF 722 protein sequence SEQ ID NO:2438.  
 XX  
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KW antibacterial; gene therapy.  
 XX  
 OS *Neisseria meningitidis*.  
 XX  
 PN WO9957280-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 PF 30-APR-1999; 99WO-US09346.  
 XX  
 PR 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0098994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.  
 XX  
 PA (CHIR) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 XX  
 DR WPI; 2000-062150/05.  
 DR N-PSDB; AAZ54244.  
 XX  
 XX Novel *Neisseria* polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics  
 XX  
 PS Claim 2; Page 1168; 1453pp; English.  
 XX  
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941  
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the  
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 XX  
 SQ Sequence 351 AA;

Query Match 0.5%; Score 7; DB 21; Length 351;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 NALVLKP 1164  
 |||||  
 DB 253 nalvlkp 259

RESULT 150  
 AAY75483  
 ID AAY75483 standard; Protein; 351 AA.  
 XX  
 AC AAY75483;  
 XX  
 DT 21-MAR-2000 (first entry)  
 XX  
 DE *Neisseria meningitidis* ORF 722 protein sequence SEQ ID NO:2440.  
 XX  
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KW antibacterial; gene therapy.  
 XX  
 OS *Neisseria meningitidis*.  
 XX  
 PN WO9957280-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 PF 30-APR-1999; 99WO-US09346.  
 XX  
 PR 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0098994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103796.  
 PR 25-FEB-1999; 99US-0121528.

```

XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
PI
XX WPI: 2000-062150/05.
DR N-PSDB; AA254245.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
PS Claim 2; Page 1169; 1453pp; English.
XX
CC AA253015 to AA254536, AA254577 to AA254615, and AA274253 to AA275941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
XX Sequence 351 AA;
SQ

```

Query Match 0.5%; Score 7; DB 21; Length 351;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1158 NALVLKP 1164
Db 253 nalvlkp 259

```

RESULT 151  
AAG40467  
ID AAG40467 standard; Protein; 354 AA.  
XX AC  
XX AAG40467;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 50214.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.

```

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

```

```
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159299.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21: Length 354;
Best Local Similarity 100.0%; Fred. No. 3.4e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 NTDDIKA 789
Db 48 ntddika 54
|||||

RESULT 152
AAG06503
ID AAG06503 standard; Protein; 363 AA.
XX
AC AAG06503;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3300.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
```

PR	30-APR-1999;	9905-0132407;
PR	04-MAY-1999;	9905-0132484
PR	05-MAY-1999;	9905-0132485
PR	06-MAY-1999;	9905-0132486
PR	06-MAY-1999;	9905-0132487
PR	07-MAY-1999;	9905-0132488
PR	11-MAY-1999;	9905-0132489
PR	14-MAY-1999;	9905-0132490
PR	14-MAY-1999;	9905-0134218
PR	14-MAY-1999;	9905-0134219
PR	14-MAY-1999;	9905-0134221
PR	14-MAY-1999;	9905-0134370
PR	18-MAY-1999;	9905-0134371
PR	19-MAY-1999;	9905-0134768
PR	20-MAY-1999;	9905-0134941
PR	21-MAY-1999;	9905-0135124
PR	22-MAY-1999;	9905-0135353
PR	24-MAY-1999;	9905-0135629
PR	25-MAY-1999;	9905-0136021
PR	27-MAY-1999;	9905-0136392
PR	28-MAY-1999;	9905-0136788
PR	01-JUN-1999;	9905-0137222
PR	03-JUN-1999;	9905-0137528
PR	04-JUN-1999;	9905-0137502
PR	07-JUN-1999;	9905-0137724
PR	08-JUN-1999;	9905-0138094
PR	10-JUN-1999;	9905-0138540
PR	10-JUN-1999;	9905-0138847
PR	14-JUN-1999;	9905-0139119
PR	16-JUN-1999;	9905-0139452
PR	17-JUN-1999;	9905-0139453
PR	17-JUN-1999;	9905-0139492
PR	18-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139455
PR	18-JUN-1999;	9905-0139456
PR	18-JUN-1999;	9905-0139457
PR	18-JUN-1999;	9905-0139458
PR	18-JUN-1999;	9905-0139459
PR	18-JUN-1999;	9905-0139460
PR	18-JUN-1999;	9905-0139461
PR	18-JUN-1999;	9905-0139462
PR	18-JUN-1999;	9905-0139463
PR	18-JUN-1999;	9905-0139750
PR	18-JUN-1999;	9905-0139763
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0139899
PR	23-JUN-1999;	9905-0140353
PR	23-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140695
PR	28-JUN-1999;	9905-0140823
PR	29-JUN-1999;	9905-0140991
PR	30-JUN-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	01-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142350
PR	06-JUL-1999;	9905-0142595
PR	08-JUL-1999;	9905-0142930
PR	09-JUL-1999;	9905-0142930
PR	13-JUL-1999;	9905-0142977
PR	14-JUL-1999;	9905-0143542
PR	14-JUL-1999;	9905-0143624
PR	15-JUL-1999;	9905-0144005
PR	16-JUL-1999;	9905-0144085
PR	16-JUL-1999;	9905-0144086
PR	19-JUL-1999;	9905-0144325
PR	19-JUL-1999;	9905-0144331
PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144333
PR	19-JUL-1999;	9905-0144334
PR	20-JUL-1999;	9905-0144352
PR	20-JUL-1999;	9905-0144352
PR	20-JUL-1999;	9905-0144632
PR	21-JUL-1999;	9905-0144884
PR	21-JUL-1999;	9905-0144814
PR	21-JUL-1999;	9905-0145086

PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145145;
PR	23-JUL-1999;	9905-0145224;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147200;
PR	04-AUG-1999;	9905-0147300;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	30-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155139;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0156559;
PR	24-SEP-1999;	9905-0156458;
PR	24-SEP-1999;	9905-0156596;
PR	04-OCT-1999;	9905-0157117;
PR	03-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158029;
PR	08-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0159293;
PR	13-OCT-1999;	9905-0159294;
PR	13-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	14-OCT-1999;	9905-0159638;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;



```

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 23-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      0.5%; Score 7; DB 21; Length 363;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 456 VDGPLRV 462
DB 131 vdgplrv 137

RESULT 153
AAG40466
ID AAG40466 standard; Protein; 365 AA.
XX
AC AAG40466;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 50213.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125789.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132046.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

```

PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 08-AUG-1999; 99US-0147433.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0156659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159233.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 0.5%; Score 7; DB 21; Length 365;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 783 NTDDIKA 789  
Db 59 ntddika 65  
|||||  
  
RESULT 154  
AAP70457  
ID AAP70457 standard; protein; 377 AA.  
XX  
AC AAP70457;  
XX  
DT 13-FEB-1991 (first entry)  
XX  
DE Sequence of gpc encoded by segment of Xanthomonas campestris DNA  
DE that contains a gene cluster that directs Xanthan biosynthesis.  
XX  
KW Thickening agent; oil recovery; drilling fluid.  
XX  
OS Xanthomonas campestris.  
XX  
PN W08705938-A.  
XX  
PD 08-OCT-1987.  
XX  
PF 24-MAR-1987; 87WO-US00604.  
XX  
PR 23-MAR-1987; 87US-0029530.  
XX  
PA (GETT-) GETTY SCI DEV CO.  
XX  
PI Capage MA, Doherty DH, Betlach MR, Vanderslice RW;  
XX  
DR WPI; 1987-291651/41.  
XX  
PT Recombinant DNA prodn. of xanthan gum or its variants - by  
PT transforming host cells with vector contg. DNA coding for enzymes  
PT involved in polysaccharide synthesis  
XX  
PS Example; Fig 12; 149pp; English.  
XX  
CC Virtually all of the segment of Xanthomonas campestris DNA that  
CC contains a gene cluster that directs xanthan biosynthesis (AAN70753),  
CC codes for protein products. Each gene is designated by a letter (see  
CC Fig 11) and its protein product is designated by that letter  
CC preceded by 'gp' (AAP70455-67).  
XX  
SQ Sequence 377 AA;

Query Match 0.5%; Score 7; DB 8; Length 377;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 976 ILNSRLV 982  
Db 143 ilnsrlv 149  
|||||

RESULT 155

AAW60766  
ID AAW60766 standard; Protein; 378 AA.

XX AAW60766;

XX 28-SEP-1998 (first entry)

DE Murine Lunatic Fringe protein.

KW Lunatic Fringe; mouse; Notch receptor; Notch ligand;  
KW cell differentiation; atherosclerosis; angioplasty;  
KW cardiovascular disease; angiogenesis; cancer; therapy.

XX Mus sp.

XX W09817793-A1.

XX 30-APR-1998.

XX 20-OCT-1997; 97WO-CA00775.

XX 21-OCT-1996; 96US-0028398.

XX (HSCR-) HSC RES & DEV LP.

PI Cohen BL, Egan SE, Lipshitz HD, Phillips RA;

XX WPI; 1998-261491/23.

DR N-PSDB; AAV36217.

XX New isolated mammalian Fringe gene(s) : used to develop products  
PT for treating, e.g. atherosclerosis, angioplasty, cardiovascular  
PT disease, angiogenesis or cancer

XX Claim 2; Page 63-64; 89pp; English.

CC This polypeptide sequence comprises novel murine Lunatic Fringe  
CC protein. The amino acid sequence was deduced from a cDNA clone  
CC (see AAV36217) isolated from an embryo cDNA library. 3 Genes (see  
CC AAV36217-19) encoding novel murine Lunatic Fringe, Manic Fringe (see  
CC AAV60767) and Radical Fringe (see AAW60768) have been identified.  
CC Undifferentiated mammalian cells appear to express Lunatic Fringe,  
CC but not Manic or Radical Fringe. During differentiation, there is  
CC a switch over to expression of Manic and Radical and a cessation of  
CC expression of Lunatic. The Fringe proteins control or modulate  
CC activation of the Notch receptor by Notch ligands. The invention  
CC provides a method for promoting cell differentiation by suppressing  
CC expression of Lunatic Fringe protein in the cell and/or promoting  
CC expression of Radical Fringe and/or Manic Fringe. A method for  
CC suppressing differentiation of a cell involves suppressing expression  
CC of Radical Fringe and/or Manic Fringe and/or promoting expression of  
CC Lunatic Fringe in the cell. The Fringe system of proteins can be  
CC used to induce new cell fates at tissue boundaries, to reinforce  
CC predetermined tissue boundaries and to block Notch signalling in  
CC differentiating cells. Polypeptides and nucleic acid sequences  
CC (including antisense sequences) can be used to prepare therapeutics  
CC useful for regulating, treating or preventing symptoms related to  
CC angioplasty, atherosclerosis, cardiovascular disease or diseases  
CC related to angiogenesis, including cancer (claimed). They can also  
CC be used to regulate skin growth and differentiation (claimed).  
CC The products can also be used for detection, diagnosis and drug  
CC screening especially for compounds that modulate expression of a  
CC mammalian Fringe gene or that selectively bind to a mammalian Fr  
CC (claimed).

XX Sequence 378 AA;

Query Match

Best Local Similarity 0.5%; Score 7; DB 19; Length 378;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 LALV GAL 23  
DB 11 lalv gal 17

RESULT 156

AAG32503  
ID AAG32503 standard; Protein; 380 AA.

XX AAG32503;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 39221.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 08-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0128714.

XX 19-APR-1999; 99US-0129845.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 18-MAY-1999; 99US-0134370.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 25-MAY-1999; 99US-0135629.

XX 27-MAY-1999; 99US-0136021.

XX 28-MAY-1999; 99US-0136392.

XX 01-JUN-1999; 99US-0136782.

XX 03-JUN-1999; 99US-0137222.

XX 04-JUN-1999; 99US-0137528.

XX 07-JUN-1999; 99US-0137502.

XX 08-JUN-1999; 99US-0137724.

XX 10-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 14-JUN-1999; 99US-0138847.

XX 16-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 380;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1168 VSYNHLG 1174  
Db 203 vsynhlg 209

RESULT 157

AAG21935  
ID AAG21935 standard; Protein; 381 AA.  
XX  
AC AAG21935;  
XX  
AC AAG21935;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24675.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
PN EF1033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0123788.  
PR 23-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 07-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 13-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-014926.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 381;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 336 NNTPSQS 342  
 |||||  
 Db 4 nntpsqs 10

RESULT 158  
 AAW98493  
 ID AAW98493 standard; Protein; 387 AA.  
 XX AC AAW98493;  
 XX AC AAW98493;  
 DT 31-MAR-1999 (first entry)  
 XX

DE H. pylori GHPO 1072 protein.  
 XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.  
 XX Helicobacter pylori.  
 OS WO9843478-A1.  
 XX 08-OCT-1998.  
 XX 01-APR-1998; 98WO-US06371.  
 XX 29-JUL-1997; 97US-0902615.  
 PR 01-APR-1997; 97US-0833457.  
 PR 24-JUN-1997; 97US-0881227.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
 PI WPI; 1998-542293/46.  
 XX N-PSDB; AAX14212.  
 DR New isolated Helicobacter polynucleotides - used to develop products  
 DR for the diagnosis, prevention and treatment of Helicobacter  
 XX infections and gastrointestinal diseases  
 PT Claim 8; Page 902-904; 2054pp; English.  
 XX This sequence represents a Helicobacter pylori GHPO protein of the  
 CC invention. The polypeptides can be used for preventing or treating  
 CC Helicobacter infections, and gastroduodenal diseases associated with  
 CC these infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
 CC used for the production of antibodies. The products can also be used for  
 CC detection and diagnosis.  
 XX SQ Sequence 387 AA;

Query Match 0.5%; Score 7; DB 19; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 405 TNADGTI 411.  
 |||||  
 Db 266 tnadgti 272

RESULT 159  
 AAY17202  
 ID AAY17202 standard; Protein; 387 AA.  
 XX AC AAY17202;  
 XX AC AAY17202;  
 DT 03-AUG-1999 (first entry)  
 XX

DE H. pylori outer membrane polypeptide.  
 XX Outer membrane polypeptide; OMP; vaccine; H. pylori infection; humoral;  
 KW cellular immune response.  
 KW Helicobacter pylori.  
 XX WO9921959-A2.  
 XX 06-MAY-1999.  
 XX 28-OCT-1998; 98WO-US22883.  
 XX 17-DEC-1997; 97US-0993001.  
 PR

PR 28-OCT-1997; 97US-0959131.  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PI Alm RA, Ellis RW, Guild BC, Noonan BM, Smith D;  
 XX WPI; 1999-326698/27.  
 DR N-PSDB; AAX75821.  
 XX Cellular vaccine against *Helicobacter pylori*  
 PT Claim 7; Page 282-283; 352pp; English.  
 XX The invention relates to a vaccine for preventing or treating infections  
 CC by *Helicobacter pylori*. The vaccine contains at least one isolated  
 CC H. pylori polypeptide, or its fragments, in a carrier, where the  
 CC carrier is a *Salmonella*, *Vibrio cholerae* or *Shigella* vector containing a  
 CC nucleic acid encoding the H. pylori polypeptide. The vaccines induce  
 CC humoral and cellular immune responses. The vaccines are used to treat or  
 CC prevent infections by H. pylori. Sequences AAX75779 to AAX75837 represent  
 CC (OMPs) AAX17160 to AAX17218.  
 XX Sequence 387 AA;  
 SQ

Query Match 0.5%; Score 7; DB 20; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 405 TNADGTI 411  
 Db 266 tnaedgti 272  
 |||||

RESULT 160  
 AAG21934  
 ID AAG21934 standard; Protein; 387 AA.  
 XX  
 AC AAG21934;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 24674.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.

30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136382.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145931.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150894.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 NNTPSQS 342

Db 10 mntpsqs 16

RESULT 161

AAW98760

ID AAW98760 standard; Protein; 395 AA.

XX AC AAW98760;

XX 31-MAR-1999 (first entry)

DE H. pylori GHPO 1087 protein.

KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

OS peptic ulcer disease.

XX Helicobacter pylori.

XX WO9843478-Al.

XX 08-OCT-1998.

PF 01-APR-1998; 98WO-US06371.

PR 29-JUL-1997; 97US-0902615.

PR 01-APR-1997; 97US-0833457.

PR 24-JUN-1997; 97US-0881227.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (JNMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;

XX WPI: 1998-542293/46.

XX N-PSDB; AAX14479.

XX New Isolated Helicobacter polynucleotides - used to develop products

PT for the diagnosis, prevention and treatment of Helicobacter

PT infections and gastrointestinal diseases

XX Claim 8; Page 1627-1629; 2054pp; English.

XX This sequence represents a Helicobacter pylori GHPO protein of the

CC invention. The polypeptides can be used for preventing or treating

CC Helicobacter infections, and gastroduodenal diseases associated with

CC these infections, including acute, chronic, and atrophic gastritis, and

CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be

CC used for the production of antibodies. The products can also be used for



CC detection and diagnosis.

XX Sequence 395 AA;

SQ Query Match 0.5%; Score 7; DB 19; Length 395;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 IGKGGVN 435

Db 343 igkggyn 349

RESULT 162

AAB76632

ID AAB76632 standard; Protein; 421 AA.

XX AC AAB76632;

XX DT 11-APR-2001 (first entry)

XX Corynebacterium glutamicum MCT protein SEQ ID NO:246.

KW Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;  
 KW membrane construction and membrane transport protein; petroleum spill;  
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;  
 KW identification; microorganism; fine chemical production; transformation;  
 KW genome mapping; genetic engineering.

XX OS Corynebacterium glutamicum.

XX PN WO200100805-A2.

XX PD 04-JAN-2001.

XX PF 23-JUN-2000; 2000WO-IB00926.

XX PR 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031454.

PR 08-JUL-1999; 99DE-1031478.

PR 08-JUL-1999; 99DE-1031563.

PR 09-JUL-1999; 99DE-1032122.

PR 09-JUL-1999; 99DE-1032124.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032180.

PR 09-JUL-1999; 99DE-1032182.

PR 09-JUL-1999; 99DE-1032190.

PR 09-JUL-1999; 99DE-1032191.

PR 09-JUL-1999; 99DE-1032209.

PR 09-JUL-1999; 99DE-1032212.

PR 09-JUL-1999; 99DE-1032227.

PR 09-JUL-1999; 99DE-1032228.

PR 09-JUL-1999; 99DE-1032229.

PR 09-JUL-1999; 99DE-1032230.

PR 14-JUL-1999; 99DE-1032927.

PR 14-JUL-1999; 99DE-1033005.

PR 14-JUL-1999; 99DE-1033006.

PR 27-AUG-1999; 99DE-1040764.

PR 27-AUG-1999; 99DE-1040765.

PR 27-AUG-1999; 99DE-1040766.

PR 27-AUG-1999; 99DE-1040830.

PR 27-AUG-1999; 99DE-1040831.

PR 27-AUG-1999; 99DE-1040832.

PR 27-AUG-1999; 99DE-1040833.

PR 31-AUG-1999; 99DE-1041378.

PR 31-AUG-1999; 99DE-1041379.

PR 31-AUG-1999; 99DE-1041395.

PR 03-SEP-1999; 99DE-1042077.

PR 03-SEP-1999; 99DE-1042078.

PR 03-SEP-1999; 99DE-1042079.

PR 03-SEP-1999; 99DE-1042088.

XX

PA (BADI ) BASF AG.

XX

PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX

DR WPI; 2001-071486/08.

XX

DR N-PSDB; AAF67865.

XX

PT Corynebacterium glutamicum nucleic acids encoding membrane construction  
 PT and membrane transport proteins or their portions, useful for typing or  
 PT identifying C. glutamicum or related bacteria, and as markers for  
 PT transformation

XX

PS Claim 20; Page 514-515; 1119pp; English.

XX

CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane  
 CC construction and membrane transport (MCT) proteins given in AAB76510 to  
 CC AAB76847. The MCT nucleic acids and proteins are useful in the  
 CC identification of microorganisms which can be used to produce fine  
 CC chemicals, for modulating fine chemical production in C. glutamicum or  
 CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or  
 CC identification of C. glutamicum or related bacteria, as reference points  
 CC for mapping C. glutamicum genome, and as markers for transformation.  
 CC AAF68082 and AAF68082 represent sequencing primers which are used in an  
 CC example from the present invention.

XX

SQ Sequence 421 AA;

Query Match

Best Local Similarity 0.5%; Score 7; DB 22; Length 421;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 917 LQTLIID 923

Db 28 lqtliliid 34

RESULT 163

AAG40913

ID AAG40913 standard; Protein; 438 AA.

XX

AC AAG40913;

XX

DT 18-OCT-2000 (first entry)

XX

DE Zea mays protein fragment SEQ ID NO: 50829.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.

XX OS Zea mays subsp. mays.

XX

PN EPI033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 06-APR-1999; 99US-0128234.

PR

PR 08-APR-1999; 99US-0128714.

PR

PR 16-APR-1999; 99US-0129845.

PR

PR 19-APR-1999; 99US-0130077.

PR

PR 21-APR-1999; 99US-0130449.

PR

PR 23-APR-1999; 99US-0130510.

PR

PR	23-APR-1999;	99US-0130891.	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134221.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134370.	PR	26-JUL-1999;	99US-0145276.
PR	18-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145913.
PR	19-MAY-1999;	99US-0134941.	PR	27-JUL-1999;	99US-0145918.
PR	20-MAY-1999;	99US-0135124.	PR	27-JUL-1999;	99US-0145919.
PR	21-MAY-1999;	99US-0135353.	PR	28-JUL-1999;	99US-0145951.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146386.
PR	25-MAY-1999;	99US-0136021.	PR	02-AUG-1999;	99US-0146388.
PR	27-MAY-1999;	99US-0136392.	PR	03-AUG-1999;	99US-0146389.
PR	28-MAY-1999;	99US-0136782.	PR	03-AUG-1999;	99US-0147038.
PR	03-JUN-1999;	99US-0137222.	PR	04-AUG-1999;	99US-0147204.
PR	04-JUN-1999;	99US-0137528.	PR	04-AUG-1999;	99US-0147302.
PR	07-JUN-1999;	99US-0137502.	PR	05-AUG-1999;	99US-0147192.
PR	08-JUN-1999;	99US-0137724.	PR	05-AUG-1999;	99US-0147260.
PR	10-JUN-1999;	99US-0138094.	PR	06-AUG-1999;	99US-0147303.
PR	10-JUN-1999;	99US-0138540.	PR	06-AUG-1999;	99US-0147416.
PR	14-JUN-1999;	99US-0138847.	PR	09-AUG-1999;	99US-0147493.
PR	16-JUN-1999;	99US-0139119.	PR	09-AUG-1999;	99US-0147935.
PR	16-JUN-1999;	99US-0139452.	PR	10-AUG-1999;	99US-0148171.
PR	16-JUN-1999;	99US-0139453.	PR	11-AUG-1999;	99US-0148319.
PR	17-JUN-1999;	99US-0139492.	PR	12-AUG-1999;	99US-0148341.
PR	18-JUN-1999;	99US-0139454.	PR	13-AUG-1999;	99US-0148565.
PR	18-JUN-1999;	99US-0139456.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139457.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139458.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139459.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139461.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139462.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139463.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139463.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139750.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139753.	PR	26-AUG-1999;	99US-0150884.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151065.
PR	22-JUN-1999;	99US-0139899.	PR	27-AUG-1999;	99US-0151066.
PR	23-JUN-1999;	99US-0140353.	PR	27-AUG-1999;	99US-0151080.
PR	23-JUN-1999;	99US-0140354.	PR	27-AUG-1999;	99US-0151080.
PR	24-JUN-1999;	99US-0140695.	PR	30-AUG-1999;	99US-0151303.
PR	28-JUN-1999;	99US-0140823.	PR	31-AUG-1999;	99US-0151438.
PR	28-JUN-1999;	99US-0140921.	PR	01-SEP-1999;	99US-0151930.
PR	30-JUN-1999;	99US-0141287.	PR	07-SEP-1999;	99US-0152363.
PR	01-JUL-1999;	99US-0141842.	PR	10-SEP-1999;	99US-0153070.
PR	01-JUL-1999;	99US-0142154.	PR	13-SEP-1999;	99

PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 438;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1158 NALVLKP 1164  
 |||||  
 Db 129 nalvlkp 135

## RESULT 164

AAB21037  
 ID AAB21037 standard; Protein; 443 AA.

XX AC AAB21037;

DT 19-DEC-2000 (first entry)

XX Human nucleic acid-binding protein, NuABP-41.

XX Human nucleic acid-binding protein; NuABP; agonist; antagonist; EST;  
 KW expressed sequence tag; drug screening; recombinant expression; antibody;  
 KW reproductive disorder; infertility; immunological disorder;  
 KW neurological disorder; cell proliferative disorder; cancer; tumour.

XX Homo sapiens.

OS WO200044900-A2.

PN 03-AUG-2000.

PD 28-JAN-2000; 2000WO-US02237.

PF 29-JAN-1999; 99US-0117904.

PR 29-JAN-1999; 99US-0117905.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Lal P, Hillman JL, Yue H, Azimzal Y, Lu AMD, Baughn MR;  
 PI Tran B, Shih LL, Au-Young JL;  
 PI WPI: 2000-499332/44.

DR N-PSDB; AAA72422.

XX Novel nucleic acid binding proteins, used to identify agonists and  
 PT antagonists of them, for the treatment of reproductive, immunological,  
 PT neurological and cell proliferative disorders including cancer.

XX Disclosure; Page 131-132; 180pp; English.

XX Sequences AAB20997-B21051 represent novel human nucleic acid-binding  
 CC proteins (NuABPs) which are encoded by the cDNA sequences

CC AAB72382-A72436. The cDNAs were produced by extension from an appropriate  
 CC EST (expressed sequence tag) using primers designed using the EST. The  
 CC invention also relates to expression constructs, host cells and  
 CC transgenic organisms comprising a human NuABP nucleic acid, recombinant  
 CC production of the human NuABPs, and antibodies against the human NuABPs,  
 CC and also to methods of screening modulators of human NuABP activity or  
 CC expression. The human NuABPs, and their agonists and antagonists are used  
 CC to treat diseases associated with overexpression or underexpression of  
 CC functional NuABPs. Human NuABP proteins and nucleotides, and NuABP  
 CC agonists and antagonists can be used to diagnose, treat and prevent  
 CC reproductive, immunological, neurological and cell proliferative  
 CC disorders. Reproductive disorders that may be treated using compositions  
 CC of the invention include infertility, endometriosis, disruptions of the  
 CC menstrual cycle and disruptions of spermatogenesis. Immunological  
 CC disorders that may be treated include AIDS, allergies, and autoimmune  
 CC disorders such as multiple sclerosis, rheumatoid arthritis, diabetes and  
 CC systemic lupus erythematosus. Neurological disorders that may be treated  
 CC include epilepsy, neurodegenerative conditions such as Alzheimer's  
 CC disease and Parkinson's disease, and mental disorders such as  
 CC Creutzfeldt-Jakob disease, and mental disorders such as schizophrenia.  
 CC Cell proliferative disorders that may be treated include a wide variety  
 CC of cancers, and also arteriosclerosis, atherosclerosis, cirrhosis and  
 CC psoriasis.

XX SQ Sequence 443 AA;

Query Match 0.5%; Score 7; DB 21; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LKQAEAA 71

Db 131 lkqaea 137  
 |||||

## RESULT 165

AAY87780

ID AAY87780 standard; Protein; 443 AA.

XX AC AAY87780;

DT 24-AUG-2000 (first entry)

XX Human ADA2 protein.

XX Human; ADA2; cytostatic; gene therapy; treatment; cancer.

XX Homo sapiens.

XX US6054289-A.

XX 25-APR-2000.

XX 30-AUG-1996; 96US-0705771.

XX 30-AUG-1995; 95US-0002993.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA;

XX WPI: 2000-338491/29.

XX N-PSDB; AAA39471.

XX New polynucleotide encoding human AD2 is useful for treating cancer and  
 PT for isolating cDNAs and genes having similar biological activity -

XX Disclosure; Column 53-56; 54pp; English.

XX This invention describes a novel polynucleotide (I) encoding human ADA2.  
 CC The products of the invention have cytostatic activity and can be used  
 CC for gene therapy. (I) is useful for treating cancer; as primers and

CC probes for isolating full length cDNA and genes having similar  
CC biological activity. This sequence represents the human ADA2 protein  
CC described in the method of the invention.

XX SQ Sequence 443 AA;

Query Match 0.5%; Score 7; DB 21; Length 443;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 65 LKQAEAA 71  
|||||||  
Db 131 lkqaeaa 137

RESULT 166  
AAG09862  
ID AAG09862 standard; Protein; 446 AA.

XX AC AAG09862;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7958.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX RW hydriation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-01471192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149802.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158039.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160787.  
PR 21-OCT-1999; 99US-0160788.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.  
PR 05-AUG-1999; 99US-01471192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149802.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158039.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160787.  
PR 21-OCT-1999; 99US-0160788.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.  
Query Match 0.5%; Score 7; DB 21; Length 446;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1168 VSYNHLG 1174  
Db 269 vsynhlg 275  
RESULT 167  
AAG32502  
ID AAG32502 standard; Protein; 446 AA.  
XX  
AC AAG32502;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 39220.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.

```
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142053.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144003.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 446;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1168 VSYNHLG 1174
Db 269 vsynhlg 275
```

RESULT 168  
 AAY31654  
 ID AAY31654 standard; Protein; 479 AA.  
 XX AC  
 XX AAY31654;  
 XX 09-NOV-1999 (first entry)  
 XX DE HLA-DR2 alpha-Fos-IgG fusion protein.  
 XX Major histocompatibility complex Class II; MHC; binding domain;  
 KW HLA-DR2; leucine zipper; Fos; IgG; Fc; immunoglobulin; antibody;  
 KW fusion protein; multiple sclerosis; rheumatoid arthritis;  
 KW graft rejection; allergy; autoimmune disease; pemphigus vulgaris;  
 KW systemic lupus erythematosus; T lymphocyte; T cell; diagnosis;  
 KW therapy; adoptive immunotherapy.  
 XX Chimeric - Homo sapiens.  
 OS Chimeric - Saccharomyces cerevisiae.  
 OS Chimeric - synthetic.  
 XX FH Key Location/Qualifiers  
 FT Peptide 1..5  
 FT Protein /note= "alpha-mating factor secretion signal"  
 FT Domain /note= "mature protein"  
 FT Peptide 8..198  
 FT Peptide /note= "DRA\*0101 extracellular domain"  
 FT Domain 206..245  
 FT Domain /note= "linker"  
 FT Domain 206..245  
 FT Domain /note= "Fos leucine zipper domain"  
 FT Domain 246..479  
 FT Domain /note= "IgG"  
 XX W09942597-A1.  
 XX 26-AUG-1999.  
 XX 19-FEB-1999; 99WO-US03603.  
 XX 19-FEB-1998; 98US-0075351.  
 XX (HARD ) HARVARD COLLEGE.  
 XX Strominger JL, Wucherpennig KW;  
 XX WPI; 1999-527481/44.  
 XX N-PSDB; AAX87813.  
 XX New MHC Class II binding domain fusion proteins and conjugates -  
 XX used for, e.g. treating allergic and autoimmune diseases or  
 XX detecting, isolating, activating or killing specific T cells  
 XX Example 7; Page 102-103; 113pp; English.  
 XX The present sequence represents a divalent HLA-DR2 MHC binding  
 XX domain fusion protein comprising an alpha-mating factor secretion  
 XX signal, the extracellular domain of the HLA-DR2 alpha chain  
 XX (residues 1-191 of DRA\*0101), a 7-amino acid linker, the 40-amino  
 XX acid leucine zipper dimerization domain of Fos, and the Fc portion  
 XX of IgG2a. The DR-alpha-Fc chain corresponds to an antibody heavy  
 XX chain. The invention provides new monovalent, multivalent and  
 XX multimeric MHC Class II binding domain fusion proteins and  
 XX conjugates comprising at least a binding domain of an MHC Class II  
 XX alpha or beta chain and a dimerization domain. The MHC fusion  
 XX proteins and conjugates can be used: for detecting and isolating T  
 XX cells having a defined MHC/peptide complex specificity (claimed);  
 XX to confer to a subject adoptive immunity to a defined MHC/peptide  
 XX complex (claimed); to stimulate or activate T cells reactive to a  
 XX defined MHC/peptide complex (claimed); for selective killing of T

CC cells reactive to a defined MHC complex (claimed); to tolerize a  
 CC subject to a defined MHC/peptide complex (claimed); to treat  
 CC allergic and autoimmune diseases, e.g. multiple sclerosis,  
 CC rheumatoid arthritis, pemphigus vulgaris, and systemic lupus  
 CC erythematosus; and to prevent organ or tissue transplant rejection.  
 CC The DR2-IgG design was chosen to increase the affinity for the T  
 CC cell receptor by increasing valency, and to attach an effector  
 CC domain, the Fc region of IgG2a. Complement fixation may result in  
 CC the lysis of target T cells following binding of DR2-IgG molecules  
 CC to the T cell receptor. DR2-IgG molecules may therefore be useful  
 CC for the selective depletion of autoaggressive T cells.  
 XX SQ Sequence 479 AA;  
 Query Match 0.5%; Score 7; DB 20; Length 479;  
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1109 FEAGQAL 1115  
 DB 59 feaggal 65  
 RESULT 169  
 AAG25739  
 ID AAG25739 standard; Protein; 495 AA.  
 XX AC AAG25739;  
 XX 17-OCT-2000 (first entry)  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 29927.  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX Arabidopsis thaliana.  
 XX EPI033405-A2.  
 XX 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-0301439.  
 XX 25-FEB-1999; 99US-0121825.  
 XX 05-MAR-1999; 99US-0123180.  
 XX 09-MAR-1999; 99US-0123548.  
 XX 23-MAR-1999; 99US-0125788.  
 XX 25-MAR-1999; 99US-0126264.  
 XX 29-MAR-1999; 99US-0126785.  
 XX 01-APR-1999; 99US-0127462.  
 XX 06-APR-1999; 99US-0128234.  
 XX 08-APR-1999; 99US-0128714.  
 XX 16-APR-1999; 99US-0129845.  
 XX 19-APR-1999; 99US-0130077.  
 XX 21-APR-1999; 99US-0130449.  
 XX 23-APR-1999; 99US-0130510.  
 XX 28-APR-1999; 99US-0130891.  
 XX 30-APR-1999; 99US-0131449.  
 XX 30-APR-1999; 99US-0132048.  
 XX 04-MAY-1999; 99US-0132407.  
 XX 05-MAY-1999; 99US-0132484.  
 XX 06-MAY-1999; 99US-0132485.  
 XX 06-MAY-1999; 99US-0132486.  
 XX 07-MAY-1999; 99US-0132487.  
 XX 11-MAY-1999; 99US-0132863.  
 XX 14-MAY-1999; 99US-0134256.  
 XX 14-MAY-1999; 99US-0134218.  
 XX 14-MAY-1999; 99US-0134219.  
 XX 14-MAY-1999; 99US-0134221.  
 XX 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 22-MAY-1999; 99US-0135529.  
PR 23-MAY-1999; 99US-0136021.  
PR 24-MAY-1999; 99US-0136392.  
PR 25-MAY-1999; 99US-0136782.  
PR 26-MAY-1999; 99US-0137222.  
PR 01-JUN-1999; 99US-0137528.  
PR 02-JUN-1999; 99US-0137528.  
PR 03-JUN-1999; 99US-0137502.  
PR 04-JUN-1999; 99US-0137724.  
PR 05-JUN-1999; 99US-0138094.  
PR 06-JUN-1999; 99US-0138540.  
PR 07-JUN-1999; 99US-0138847.  
PR 08-JUN-1999; 99US-0139119.  
PR 09-JUN-1999; 99US-0139452.  
PR 10-JUN-1999; 99US-0139453.  
PR 11-JUN-1999; 99US-0139492.  
PR 12-JUN-1999; 99US-0139454.  
PR 13-JUN-1999; 99US-0139455.  
PR 14-JUN-1999; 99US-0139456.  
PR 15-JUN-1999; 99US-0139457.  
PR 16-JUN-1999; 99US-0139458.  
PR 17-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 19-JUN-1999; 99US-0139461.  
PR 20-JUN-1999; 99US-0139462.  
PR 21-JUN-1999; 99US-0139463.  
PR 22-JUN-1999; 99US-0139750.  
PR 23-JUN-1999; 99US-0139763.  
PR 24-JUN-1999; 99US-0139817.  
PR 25-JUN-1999; 99US-0139899.  
PR 26-JUN-1999; 99US-0140353.  
PR 27-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 02-JUL-1999; 99US-0141842.  
PR 03-JUL-1999; 99US-0142154.  
PR 04-JUL-1999; 99US-0142055.  
PR 05-JUL-1999; 99US-0142390.  
PR 06-JUL-1999; 99US-0142803.  
PR 07-JUL-1999; 99US-0142920.  
PR 08-JUL-1999; 99US-0142977.  
PR 09-JUL-1999; 99US-0143542.  
PR 10-JUL-1999; 99US-0143624.  
PR 11-JUL-1999; 99US-0144005.  
PR 12-JUL-1999; 99US-0144085.  
PR 13-JUL-1999; 99US-0144086.  
PR 14-JUL-1999; 99US-0144325.  
PR 15-JUL-1999; 99US-0144331.  
PR 16-JUL-1999; 99US-0144332.  
PR 17-JUL-1999; 99US-0144333.  
PR 18-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 21-JUL-1999; 99US-0144632.  
PR 22-JUL-1999; 99US-0144814.  
PR 23-JUL-1999; 99US-0145086.  
PR 24-JUL-1999; 99US-0145088.  
PR 25-JUL-1999; 99US-0145085.  
PR 26-JUL-1999; 99US-0145087.  
PR 27-JUL-1999; 99US-0145089.  
PR 28-JUL-1999; 99US-0145192.  
PR 29-JUL-1999; 99US-0145145.  
PR 30-JUL-1999; 99US-0145218.  
PR 31-JUL-1999; 99US-0145224.  
PR 01-AUG-1999; 99US-0145276.  
PR 02-AUG-1999; 99US-0145913.  
PR 03-AUG-1999; 99US-0145918.  
PR 04-AUG-1999; 99US-0145919.  
PR 05-AUG-1999; 99US-0145951.  
PR 06-AUG-1999; 99US-0146386.  
PR 07-AUG-1999; 99US-0146388.  
PR 08-AUG-1999; 99US-0146389.  
PR 09-AUG-1999; 99US-0147038.  
PR 10-AUG-1999; 99US-0147204.  
PR 11-AUG-1999; 99US-0147302.  
PR 12-AUG-1999; 99US-0147192.  
PR 13-AUG-1999; 99US-0147260.  
PR 14-AUG-1999; 99US-0147303.  
PR 15-AUG-1999; 99US-0147416.  
PR 16-AUG-1999; 99US-0147493.  
PR 17-AUG-1999; 99US-0147935.  
PR 18-AUG-1999; 99US-0148171.  
PR 19-AUG-1999; 99US-0148319.  
PR 20-AUG-1999; 99US-0148341.  
PR 21-AUG-1999; 99US-0148565.  
PR 22-AUG-1999; 99US-0148684.  
PR 23-AUG-1999; 99US-0149368.  
PR 24-AUG-1999; 99US-0149175.  
PR 25-AUG-1999; 99US-0149426.  
PR 26-AUG-1999; 99US-0149722.  
PR 27-AUG-1999; 99US-0149723.  
PR 28-AUG-1999; 99US-0149829.  
PR 29-AUG-1999; 99US-0149902.  
PR 30-AUG-1999; 99US-0149930.  
PR 31-AUG-1999; 99US-0150566.  
PR 01-SEP-1999; 99US-0150884.  
PR 02-SEP-1999; 99US-0151065.  
PR 03-SEP-1999; 99US-0151066.  
PR 04-SEP-1999; 99US-0151080.  
PR 05-SEP-1999; 99US-0151303.  
PR 06-SEP-1999; 99US-0151438.  
PR 07-SEP-1999; 99US-0151930.  
PR 08-SEP-1999; 99US-0152263.  
PR 09-SEP-1999; 99US-0153070.  
PR 10-SEP-1999; 99US-0153758.  
PR 11-SEP-1999; 99US-0154018.  
PR 12-SEP-1999; 99US-0154039.  
PR 13-SEP-1999; 99US-0154779.  
PR 14-SEP-1999; 99US-0155139.  
PR 15-SEP-1999; 99US-0155486.  
PR 16-SEP-1999; 99US-0156559.  
PR 17-SEP-1999; 99US-0156458.  
PR 18-SEP-1999; 99US-0156596.  
PR 19-SEP-1999; 99US-0157117.  
PR 20-SEP-1999; 99US-0157753.  
PR 21-SEP-1999; 99US-0157865.  
PR 22-SEP-1999; 99US-0158029.  
PR 23-SEP-1999; 99US-0158232.  
PR 24-SEP-1999; 99US-0158369.  
PR 25-SEP-1999; 99US-0159293.  
PR 26-SEP-1999; 99US-0159294.  
PR 27-SEP-1999; 99US-0159295.  
PR 28-SEP-1999; 99US-0159329.  
PR 29-SEP-1999; 99US-0159330.  
PR 30-SEP-1999; 99US-0159331.  
PR 01-OCT-1999; 99US-0159337.  
PR 02-OCT-1999; 99US-0159637.  
PR 03-OCT-1999; 99US-0159638.  
PR 04-OCT-1999; 99US-0159584.  
PR 05-OCT-1999; 99US-0160741.  
PR 06-OCT-1999; 99US-0160767.  
PR 07-OCT-1999; 99US-0160768.  
PR 08-OCT-1999; 99US-0160770.  
PR 09-OCT-1999; 99US-0160814.  
PR 10-OCT-1999; 99US-0160815.  
PR 11-OCT-1999; 99US-0160980.  
PR 12-OCT-1999; 99US-0160981.  
PR 13-OCT-1999; 99US-0160989.  
PR 14-OCT-1999; 99US-0161404.  
PR 15-OCT-1999; 99US-0161405.  
PR 16-OCT-1999; 99US-0161406.  
PR 17-OCT-1999; 99US-0161406.



```
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 28-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 495;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 522 GGENIJD 528
Db 429 ggnfntld 435

RESULT 170
AAG37947
ID AAG37947 standard; Protein; 495 AA.
XX AC AAG37947;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 46743.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135628.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

105 MAY-1999; 99US-0136782.
105 JUN-1999; 99US-0137222.
105 JUN-1999; 99US-0137528.
105 JUN-1999; 99US-0137502.
105 JUN-1999; 99US-0137724.
105 JUN-1999; 99US-0138094.
105 JUN-1999; 99US-0138540.
105 JUN-1999; 99US-0138847.
105 JUN-1999; 99US-0139119.
105 JUN-1999; 99US-0139452.
105 JUN-1999; 99US-0139453.
105 JUN-1999; 99US-0139492.
105 JUN-1999; 99US-0139454.
105 JUN-1999; 99US-0139455.
105 JUN-1999; 99US-0139456.
105 JUN-1999; 99US-0139457.
105 JUN-1999; 99US-0139458.
105 JUN-1999; 99US-0139459.
105 JUN-1999; 99US-0139460.
105 JUN-1999; 99US-0139461.
105 JUN-1999; 99US-0139462.
105 JUN-1999; 99US-0139463.
105 JUN-1999; 99US-0139750.
105 JUN-1999; 99US-0139763.
105 JUN-1999; 99US-0139817.
105 JUN-1999; 99US-0139899.
105 JUN-1999; 99US-0140353.
105 JUN-1999; 99US-0140354.
105 JUN-1999; 99US-0140695.
105 JUN-1999; 99US-0140823.
105 JUN-1999; 99US-0140991.
105 JUN-1999; 99US-0141287.
105 JUN-1999; 99US-0141842.
105 JUN-1999; 99US-0142154.
105 JUN-1999; 99US-0142055.
105 JUN-1999; 99US-0142390.
105 JUN-1999; 99US-0142803.
105 JUN-1999; 99US-0142920.
105 JUN-1999; 99US-0142977.
105 JUN-1999; 99US-0143624.
105 JUN-1999; 99US-0144005.
105 JUN-1999; 99US-0144085.
105 JUN-1999; 99US-0144086.
105 JUN-1999; 99US-0144325.
105 JUN-1999; 99US-0144331.
105 JUN-1999; 99US-0144332.
105 JUN-1999; 99US-0144333.
105 JUN-1999; 99US-0144334.
105 JUN-1999; 99US-0144335.
105 JUN-1999; 99US-0144352.
105 JUN-1999; 99US-0144632.
105 JUN-1999; 99US-0144884.
105 JUN-1999; 99US-0144814.
105 JUN-1999; 99US-0145086.
105 JUN-1999; 99US-0145088.
105 JUN-1999; 99US-0145085.
105 JUN-1999; 99US-0145087.
105 JUN-1999; 99US-0145089.
105 JUN-1999; 99US-0145192.
105 JUN-1999; 99US-0145145.
105 JUN-1999; 99US-0145218.
105 JUN-1999; 99US-0145224.
105 JUN-1999; 99US-0145276.
105 JUN-1999; 99US-0145913.
105 JUN-1999; 99US-0145918.
105 JUN-1999; 99US-0145919.
105 JUN-1999; 99US-0145951.
105 JUN-1999; 99US-0145952.
105 JUN-1999; 99US-0146386.
105 JUN-1999; 99US-0146388.
105 JUN-1999; 99US-0146389.
105 JUN-1999; 99US-0147038.
105 JUN-1999; 99US-0147204.
```

PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 495;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 522 GGFNTLD 528  
Db 429 ggntld 435  
|||||

RESULT 171  
AAG39756  
ID AAG39756 standard; Protein; 495 AA.  
XX  
AC AAG39756;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49239.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX  
XX  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0128264.  
PR 29-MAR-1999; 99US-0128785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139453.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147302.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158023.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 495;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 522 GGFNTLD 528

|||||

Db 429 ggntld 435

RESULT 172  
AAG40912  
ID AAG40912 standard; Protein: 498 AA.  
XX AC  
XX AAG40912;  
XX DT  
XX 18-OCT-2000 (first entry)  
XX DE  
XX Zea mays protein fragment SEQ ID NO: 50828.  
XX KW  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX OS  
XX Zea mays subsp. mays.  
XX PN  
XX EP1033405-A2.  
XX PD  
XX 06-SEP-2000.  
XX PF  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143342.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0156559.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 498;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 NALVLKP 1164  
 |||||  
 Db 189 nalvlkp 195

RESULT 173  
 AARI5354  
 ID AARI5354 standard; Protein; 510 AA.  
 XX

AC AARI5354;  
 XX 11-MAR-1992 (first entry)  
 DT  
 XX Protein deduced from B.thuringiensis ori43 sequence.  
 DE  
 XX origin of replication; 43mda plasmid; AT-rich region.  
 KW  
 XX Bacillus thuringiensis var. kurstaki strain HD-263.  
 OS  
 XX WO9118102-A.  
 PN  
 XX 28-NOV-1991.  
 PD  
 XX 14-MAY-1991; 91WO-US03360.  
 PF  
 XX 15-MAY-1990; 90US-0523671.  
 PR  
 XX (ECOG-) ECOGEN INC.  
 PA  
 XX Baum JA;  
 PI  
 XX WPI; 1991-369254/50.  
 DR  
 XX N-PSDB; AAQ15171.  
 DX  
 XX Shuttle vector for recombinant Bacillus strain development -  
 PT useful for inserting cloned insecticidal Bt toxin genes and  
 PT config. no. DNA from non-Bt biological sources  
 XX  
 PS Disclosure; Fig 7; 94pp; English.  
 XX  
 CC The ori43 sequence was isolated from plasmid pEG599, a 43mda  
 CC plasmid containing a B.thuringiensis ori. The protein deduced from  
 CC the ORF has a rel. molecular mass of 58,320.  
 CC See also AAQ15169 and AAQ15170.  
 XX  
 SQ Sequence 510 AA;  
 Query Match 0.5%; Score 7; DB 12; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 558 LIVKTING 564  
 |||||  
 Db 435 livkting 441  
 RESULT 174  
 AAW79740  
 ID AAW79740 standard; Protein; 510 AA.  
 XX  
 AC AAW79740;  
 XX  
 DT 02-FEB-1999 (first entry)  
 XX  
 DE Soybean wild-type myo-inositol 1-phosphate synthase.  
 XX  
 KW Soybean; myo-inositol 1-phosphate synthase; raffinose;  
 KW stachyose; phytic acid.  
 XX  
 OS Glycine max line LR13.  
 XX  
 PN WO9845448-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 07-APR-1998; 98WO-US06822.  
 XX  
 PR 08-APR-1997; 97US-0835751.  
 XX  
 PA (DUFO) DU PONT DE NEMOURS & CO E I.  
 XX

PI Hitz WD, Sebastian SA;  
 XX WPI; 1998-568353/48.  
 DR N-PSDB; AAV62440.  
 XX  
 XX Soybean plants containing altered myo-inositol-1-phosphate gene -  
 PT useful for generating plants with altered levels of e.g. raffinose,  
 PT stachyose, phytic acid, etc  
 XX  
 XX Example 5; Page 45-47; 63pp; English.  
 PS  
 XX This is the amino acid sequence of soybean myo-inositol 1-phosphate  
 CC synthase (MI 1-PS) deduced from the coding region of an isolated  
 CC cDNA clone (see AAV62440). MI 1-PS is involved in glucose metabolism  
 CC to phytic acid, raffinose and stachyose. A mutant MI 1-PS (see  
 CC AAW79741) has been identified in soybean line LR33, a mutagenised  
 CC line of low raffinose saccharide phenotype. Sequencing revealed a  
 CC single base change mutation in the LR33 gene sequence that  
 CC resulted in a K396N substitution in the mutant protein. The  
 CC mutation results in a seed phenotype of very low raffinose  
 CC saccharide sugars, very high sucrose and low phytic acid. The  
 CC mutated nucleic acid is used to alter the raffinose saccharide,  
 CC sucrose, phytic acid and inorganic phosphate content of soybean  
 CC seeds, leading to useful soybean products, e.g. a seed phytic acid  
 CC content of less than 17 ug/g, a seed content of raffinose and  
 CC stachyose combined of less than 14.5 ug/g, and a seed sucrose  
 CC content greater than 200 ug/g.  
 XX  
 SQ Sequence 510 AA;  
 Query Match 0.5%; Score 7; DB 19; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1168 VSYNHLG 1174  
 Db 333 vsynhlg 339  
 RESULT 175  
 AAW79741  
 ID AAW79741 standard; Protein; 510 AA.  
 XX  
 AC AAW79741;  
 XX  
 XX 02-FEB-1999 (first entry)  
 DT  
 XX Soybean mutant myo-inositol 1-phosphate synthase.  
 DE  
 XX Soybean; myo-inositol 1-phosphate synthase; raffinose;  
 KW stachyose; phytic acid.  
 XX  
 XX Glycine max line LR33.  
 OS  
 XX WO9845448-A1.  
 PN  
 XX 15-OCT-1998.  
 PD  
 XX 07-APR-1998; 98WO-US06822.  
 PF  
 XX 08-APR-1997; 97US-0835751.  
 PR  
 XX (DUPO) DU PONT DE NEMOURS & CO E I.  
 PA  
 XX Hitz WD, Sebastian SA;  
 PI  
 XX WPI; 1998-568353/48.  
 DR N-PSDB; AAV62443.  
 XX  
 XX Soybean plants containing altered myo-inositol-1-phosphate gene -  
 PT useful for generating plants with altered levels of e.g. raffinose,  
 PT stachyose, phytic acid, etc

XX  
 PS Example 5; Page 49-51; 63pp; English.  
 XX  
 CC This is the amino acid sequence of a mutant soybean myo-inositol  
 CC 1-phosphate synthase (MI 1-PS) deduced from the coding region of an  
 CC isolated cDNA clone (see AAV62443). MI 1-PS is involved in glucose  
 CC metabolism to phytic acid, raffinose and stachyose. The MI 1-PS  
 CC was identified in soybean line LR33, a mutagenised line of low  
 CC raffinose saccharide phenotype. Sequencing revealed a single base  
 CC change mutation in the LR33 gene sequence that resulted in a K396N  
 CC substitution in the mutant protein compared to wild-type MI 1-PS  
 CC (see AAW79740). The mutation results in a seed phenotype of very low  
 CC raffinose saccharide sugars, very high sucrose and low phytic acid.  
 CC The mutated nucleic acid is used to alter the raffinose saccharide,  
 CC sucrose, phytic acid and inorganic phosphate content of soybean  
 CC seeds, leading to useful soybean products, e.g. a seed phytic acid  
 CC content of less than 17 ug/g, a seed content of raffinose and  
 CC stachyose combined of less than 14.5 ug/g, and a seed sucrose  
 CC content greater than 200 ug/g.  
 XX  
 SQ Sequence 510 AA;  
 Query Match 0.5%; Score 7; DB 19; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1168 VSYNHLG 1174  
 Db 333 vsynhlg 339  
 RESULT 176  
 AAW96259  
 ID AAW96259 standard; Protein; 510 AA.  
 XX  
 AC AAW96259;  
 XX  
 XX 14-JUN-1999 (first entry)  
 DT  
 XX Phytate protein.  
 DE  
 XX Phytate gene; maize; soybean; wheat; barley; canola; sunflower;  
 KW inducible; germination; flour; grain.  
 XX  
 XX Zea mays.  
 OS  
 XX WO9907211-A1.  
 PN  
 XX 18-FEB-1999.  
 PD  
 XX 11-AUG-1998; 98WO-US16702.  
 PF  
 XX 11-AUG-1997; 97US-0055323.  
 PR  
 XX (EXSE-) EXSEED GENETICS LLC.  
 PA  
 XX Chang M, Guan H, Keeling PL, Wilhelm EP;  
 PI  
 XX WPI; 1999-167128/14.  
 DR N-PSDB; AAX09006.  
 XX  
 XX Transgenic plants in which production of phytic acid production is  
 PT regulated - by the presence of a genetic construct which is  
 PT operably linked to an inducible promoter  
 XX  
 XX Claim 12; Page 51-53; 58pp; English.  
 PS  
 XX Transgenic plants containing a genetic construct comprising a  
 CC nucleic acid sequence encoding a selected gene product that  
 CC regulates the plant's production of phytic acid is used to  
 CC regulate germination of seeds which comprise a non-germinable  
 CC trait. When the nucleic acid is induced and the seed produces

CC phytic acid, the non-germinable trait is corrected or overcome  
 CC and the seed germinates. The construct of the invention is useful  
 CC in crops such as soybeans, wheat, barley, maize, canola, and  
 CC sunflowers because the construct can be used to produce seed with  
 CC a very low level of phytic acid and also the seed can only germinate  
 CC in highly controlled conditions following induction. The invention  
 CC is particularly useful in the flour and feed grains, such as, corn,  
 CC wheat, soybeans, sunflower, oats, and rye. The construct of the  
 CC invention, when present in a plant is useful to farmers because it  
 CC eliminates volunteer plants in the next season by rendering the  
 CC seed incapable of germinating and prevents farmers from saving seed  
 CC by rendering the seed incapable of being reproduced for future  
 CC years. It is also useful in the feed and milling industry because  
 CC it adds nutritional value to the feed or milled products by being  
 CC low in phytic acid content.

XX Sequence 510 AA;

Query Match 0.5%; Score 7; DB 20; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 Db 333 vsynhlg 339  
 |||||

RESULT 177

AAW97862  
 ID AAW97862 standard; Protein; 510 AA.

XX AC AAW97862;

DT 07-JUN-1999 (first entry)

DE Maize myo-inositol 1-phosphate synthase.

KW Myo-inositol 1-phosphate synthase; maize; corn; phytate;  
 KW phytic acid; transgenic plant; animal nutrition; feedstuff; food.

XX Zea mays.

XX WO9905298-A1.

PD 04-FEB-1999.

XX 17-JUL-1998; 98WO-US14657.

XX 18-MAY-1998; 98US-0085852.

XX 22-JUL-1997; 97US-0053371.

XX 28-JUL-1997; 97US-0053344.

XX 08-AUG-1997; 97US-0055226.

XX 11-AUG-1997; 97US-0055446.

XX (PION-) PIONEER HI-BRED INT INC.

XX Beach LR, Bowen BA, Martino-Catt SJ, Wang H, Wang X;

XX WPI; 1999-142948/12.

XX N-PSDB; AAX24407.

XX New polynucleotides controlling phytate metabolism in plants -

XX useful for improving the nutritional content of plants, by enhancing  
 XX levels of non-phytate phosphorus, and reducing phytate levels  
 XX Claim 1a; Page 77-78; 86pp; English.  
 XX This is the amino acid sequence of maize myo-inositol 1-phosphate  
 XX synthase, an enzyme involved in the metabolism of phytate. cDNA  
 XX (see AAX24407) encoding the enzyme was isolated from a maize embryo  
 XX (15 day post-pollination) cDNA library. Polynucleotides (see  
 XX AAX24400, AAX24403, AAX24407 and AAX24410-12) encoding maize

CC phosphatidylinositol-3-kinase (see AAW97880), myo-inositol  
 CC 1,3,4-triphosphate 5/6-kinase (see AAW97881), myo-inositol  
 CC 1-phosphate synthase and myo-inositol monophosphatase-3 (see  
 CC AAW97883), all enzymes involved in phytate metabolism, are claimed.  
 CC The invention relates to the use of such genes to reduce the  
 CC levels of phytate, and/or increase the levels of non-phytate  
 CC phosphorus, in plants used for food or feed. The genes are  
 CC especially used to improve the nutritional content of plants such  
 CC as corn and soybean. Transgenic plants, and seed produced by them,  
 CC are claimed.

XX Sequence 510 AA;

Query Match 0.5%; Score 7; DB 20; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
 Db 333 vsynhlg 339  
 |||||

RESULT 178

AAW97875

ID AAW97875 standard; Protein; 510 AA.

XX AC AAW97875;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 64107.

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135622.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.



PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 510;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
|||||||  
DB 333 vsynhl9 339

RESULT 179  
AAB48935  
ID AAB48935 standard; Protein; 510 AA.

XX AAB48935;  
XX 16-MAR-2001 (first entry)  
DT  
DE  
XX

XX Brassica napus myo-inositol 1-phosphate synthase (MIPS).

KW Myo-inositol 1-phosphate synthase; MIPS; rape; canola meal; phytate;  
KW Phytic acid; antisense inhibition; cosuppression; transgenic plant;  
KW animal feed.

OS Brassica napus.

XX WO200073473-A1.

XX 07-DEC-2000.

XX 25-MAY-2000; 2000WO-CA00612.

XX 26-MAY-1999; 99US-0136204.

XX (CANA ) NAT RES COUNCIL CANADA.

PI Georges F, Hussain AA, Keller WA;

XX WPI; 2001-061548/07.

DR N-PSDB; AAC87643.

XX A method for reducing phytate in Brassica plants, especially useful in  
PT reducing phytate in the canola meal or protein used in animal feeds,  
PT comprises employing genetic manipulation involving the myo-inositol  
PT 1-phosphate synthase gene -

PS Claim 4; Page 19-22; 31pp; English.

CC The invention relates to Brassica napus myo-inositol 1-phosphate  
CC synthase (MIPS; AAB48935), to nucleic acids encoding it (AAC87643), and  
CC to the myo-inositol 1-phosphate synthase gene promoter (AAC87646).  
CC The invention also relates to cells, particularly Brassica cells,  
CC transformed with a nucleic acid of the invention, and a plant, plant  
CC fragment or seed which has been transformed with a nucleic acid of the  
CC invention in which myo-inositol 1-phosphate synthase activity is  
CC reduced. The invention additionally relates to a novel method for  
CC reducing the phytate content in plants of the genus Brassica,  
CC particularly in Brassica napus (rape, canola) by growing a such a plant  
CC which comprises a MIPS antisense sequence and/or a MIPS cosuppression  
CC sequence. As phytic acid is the hexaphosphate derivative of  
CC myo-inositol, reduced expression of MIPS will result in reduced phytate  
CC levels in the plant. Phytate is a storage substance which does not  
CC participate in any of the essential pathways during plant growth and  
CC development, but which is an undesirable component in animal feeds,  
CC particularly those based on canola meal. The method is particularly  
CC useful for reducing phytate in canola meal, which is used as a protein  
CC source in the animal feeding industry. The method is also useful for

CC enhancing protein or oil synthesis, thereby increasing the economic value  
CC of canola. The present sequence represents Brassica napus myo-inositol  
CC 1-phosphate synthase.

XX  
SQ Sequence 510 AA;

Query Match 0.5%; Score 7; DB 22; Length 510;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
|||||||  
DB 333 vsynhl9 339

RESULT 180  
AAG09861  
ID AAG09861 standard; Protein; 511 AA.

XX AAG09861;

XX 17-OCT-2000 (first entry)  
DT

DE Arabidopsis thaliana protein fragment SEQ ID NO: 7957.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132863.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137724.  
PR 10-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 21-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 08-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 511;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174  
Db 334 vsynhlg 340  
|||||

RESULT 181  
AAG32501  
ID AAG32501 standard; Protein; 511 AA.  
XX AC AAG32501;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 39219.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145921.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147152.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160983.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 511;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1168 VSYNHLG 1174

Db 334 vsynhlg 340

RESULT 182  
AAG09860

ID AAG09860 standard; Protein; 534 AA.

XX AC AAG09860;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7956.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 16-APR-1999; 99US-0128714.

XX PR 19-APR-1999; 99US-0129845.

XX PR 21-APR-1999; 99US-0130077.

XX PR 23-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132407.

XX PR 05-MAY-1999; 99US-0132484.

XX PR 06-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132487.

XX PR 11-MAY-1999; 99US-0132863.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 18-MAY-1999; 99US-0134370.

XX PR 19-MAY-1999; 99US-0134768.

XX PR 20-MAY-1999; 99US-0134941.

XX PR 21-MAY-1999; 99US-0135124.

XX PR 21-MAY-1999; 99US-0135353.

XX PR 24-MAY-1999; 99US-0135629.

XX PR 25-MAY-1999; 99US-0136021.

XX PR 27-MAY-1999; 99US-0136392.

XX PR 28-MAY-1999; 99US-0136782.

XX PR 01-JUN-1999; 99US-0137222.

XX PR 03-JUN-1999; 99US-0137528.

XX PR 04-JUN-1999; 99US-0137502.

XX PR 07-JUN-1999; 99US-0137724.

XX PR 08-JUN-1999; 99US-0138094.

XX PR 10-JUN-1999; 99US-0138540.

XX PR 10-JUN-1999; 99US-0138847.

XX PR 14-JUN-1999; 99US-0139119.

XX PR 16-JUN-1999; 99US-0139452.

XX PR 16-JUN-1999; 99US-0139453.

XX PR 17-JUN-1999; 99US-0139452.

XX PR 18-JUN-1999; 99US-0139454.

XX PR 18-JUN-1999; 99US-0139455.

XX PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139464.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 534;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1168 VSYNHLG 1174  
|||||  
Db 357 vsynhlg 363

RESULT 183  
AAY24477  
ID AAY24477 standard; Protein; 536 AA.  
XX  
AC AAY24477;

XX DT 24-SEP-1999 (first entry)  
 XX DE Nicotiana paniculata INPS protein.  
 XX KW Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;  
 XX KW water stress; resistance.  
 XX OS Nicotiana paniculata.  
 XX FH Key Location/Qualifiers  
 XX FT Misc-difference 511  
 XX FT /label= unknown  
 XX FT /note= "encoded by the stop codon TGA"  
 XX PN JP11187879-A.  
 XX PD 13-JUL-1999.  
 XX PF 26-DEC-1997; 97JP-0359773.  
 XX PR 26-DEC-1997; 97JP-0359773.  
 XX PA (NIBS ) JAPAN TOBACCO INC.  
 XX DR WPI; 1999-451546/38.  
 XX DR N-PSDB; AAX90402.  
 XX PT New INPS gene derived from Nicotiana genus plant - useful for  
 XX PT conferring resistance to water stress to plants  
 XX PS Claim 2; Page 6-8; 8pp; Japanese.  
 XX CC The present sequence is the Nicotiana paniculata inositol monophosphate  
 XX CC synthase (INPS), designated NpINPS1. INPS can be used to confer water  
 XX CC stress resistance to a plant.  
 XX SQ Sequence 536 AA;  
 Query Match 0.5%; Score 7; DB 20; Length 536;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1168 VSYNHLG 1174  
 Db 333 vsynhlg 339  
 RESULT 184  
 ID AAW41608 standard; Protein; 539 AA.  
 XX AC AAW41608;  
 XX DT 20-APR-1998 (first entry)  
 XX DE Soybean protox-1.  
 XX KW Protoporphyrinogen oxidase-1; protox-1; promoter; soybean;  
 XX KW herbicide resistance; breeding programme; probe; gene isolation;  
 XX KW genomic mapping.  
 XX OS Glycine max.  
 XX FH Key Location/Qualifiers  
 XX FT CDS 55..1686  
 XX FT /\*tag= a  
 XX FT /product= protox-1  
 XX PN WO9732028-A1.  
 XX PD 04-SEP-1997.

XX PF 27-FEB-1997; 97WO-US03343.  
 XX PR 21-JUN-1996; 96US-0020003.  
 XX PR 28-FEB-1996; 96US-0012705.  
 XX PR 28-FEB-1996; 96US-0013612.  
 XX PA (NOVS ) NOVARTIS AG.  
 XX PI Johnson MA, Volrath SL, Ward ER;  
 XX DR WPI; 1997-489209/45.  
 XX DR N-PSDB; AAV04309.  
 XX PT DNA containing a plant proto-porphyrinogen oxidase gene promoter -  
 XX PT optionally linked to a heterologous gene, especially to express  
 XX PT herbicide-resistant enzymes, and plants containing such constructs  
 XX PS Claim 33; Pages 68-71; 114pp; English.  
 XX CC The present sequence is soybean protoporphyrinogen oxidase-1  
 XX CC (protox-1).  
 XX CC The protox-1 promoter can be used to express herbicide resistant  
 XX CC enzymes, specifically protox, i.e. a plant tissue, plant or progeny  
 XX CC containing a chimeric gene of the promoter and a heterologous  
 XX CC coding sequence. The plant can also be used in breeding programmes.  
 XX CC Also hybridising fragments of the protox coding sequence can be  
 XX CC used as probes, e.g. to isolate related genes or for genomic  
 XX CC mapping.  
 XX SQ Sequence 539 AA;  
 Query Match 0.5%; Score 7; DB 18; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 106 SSKIDGG 112  
 Db 48 sskidgg 54  
 RESULT 185  
 ID AAW41609 standard; Protein; 539 AA.  
 XX AC AAW41609;  
 XX DT 20-APR-1998 (first entry)  
 XX DE Cotton protox-1.  
 XX KW Protoporphyrinogen oxidase-1; protox-1; promoter; cotton;  
 XX KW herbicide resistance; breeding programme; probe; gene isolation;  
 XX KW genomic mapping.  
 XX OS Gossypium hirsutum.  
 XX PN WO9732028-A1.  
 XX PD 04-SEP-1997.  
 XX PF 27-FEB-1997; 97WO-US03343.  
 XX PR 21-JUN-1996; 96US-0020003.  
 XX PR 28-FEB-1996; 96US-0012705.  
 XX PR 28-FEB-1996; 96US-0013612.  
 XX PA (NOVS ) NOVARTIS AG.  
 XX PI Johnson MA, Volrath SL, Ward ER;  
 XX DR WPI; 1997-489209/45.

DR N-PSDB; AAV04313.

XX DNA containing a plant proto-porphyrinogen oxidase gene promoter -

PT optionally linked to a heterologous gene, especially to express

PT herbicide-resistant enzymes, and plants containing such constructs

XX

PS Claim 35; Pages 80-82; 114pp; English.

XX

CC The present sequence is cotton protoporphyrinogen oxidase-1

CC (protox-1).

CC The protox-1 promoter can be used to express herbicide resistant

CC enzymes, specifically protox, i.e. a plant tissue, plant or progeny

CC containing a chimeric gene of the promoter and a heterologous

CC coding sequence. The plant can also be used in breeding programmes.

CC Also hybridising fragments of the protox coding sequence can be

CC used as probes, e.g. to isolate related genes or for genomic

CC mapping.

XX

XX Sequence 539 AA;

XX

Query Match 0.5%; Score 7; DB 18; Length 539;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112

Db 48 sskidgg 54

|||||

RESULT 186

AAW25740

ID AAW25740 standard; Protein; 539 AA.

XX

AC AAW25740;

XX

DT 01-MAR-1998 (first entry)

XX

DE Cotton protoporphyrinogen oxidase (protox-1).

XX

KW Protox-1; protoporphyrinogen oxidase; inhibitor; cotton;

KW herbicide tolerance; herbicide resistance; transgenic plant.

XX

OS Gossypium hirsutum L.

XX

FH Key Location/Qualifiers

FT Misc-difference 365

FT

FT /note= "substitution of Pro-365 by another amino

FT acid, especially Ser, provides a modified

FT protox tolerant to a herbicide (Claims

FT 51-52)";

FT

FT Misc-difference 428

FT

FT /note= "substitution of Tyr-428 by another amino

FT acid, especially Cys or Arg, provides a

FT modified protox tolerant to a herbicide

FT (Claims 53-54)";

XX

PN WO9732011-A1.

XX

XX

PD 04-SEP-1997.

XX

XX

PF 27-FEB-1997; 97WO-US03313.

XX

XX

PR 21-JUN-1996; 96US-0020003.

PR

PR 28-FEB-1996; 96US-0012705.

PR

XX 28-FEB-1996; 96US-0013612.

XX

PA (NOVS ) NOVARTIS AG.

XX

XX

PI Helfetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER;

XX

XX WPI; 1997-448683/41.

DR

DR N-PSDB; AAT86123.

XX New DNA encoding plant protoporphyrinogen oxidase enzyme - and

PT herbicide resistant mutants, useful to prepare plants resistant to

PT herbicide which therefore kills undesired vegetation only

XX

PS Claim 7; Page 139-141; 196pp; English.

XX

CC This protein comprises cotton protoporphyrinogen oxidase

CC (protox-1), an enzyme that catalyses the oxidation of

CC protoporphyrinogen IX to protoporphyrin IX. Its amino acid

CC sequence was deduced from an cotedyledon Protox-1 cDNA clone (see

CC AAT86123). Sites within the cotton protox-1 gene have been

CC identified that can be mutated to encode a modified protox that is

CC resistant to protox inhibitors and hence tolerant of certain

CC herbicides. Plants, especially crop plants, may be engineered for

CC resistance to protox inhibitors via mutation of the native protox

CC gene to a resistant form, or they may be transformed with a gene

CC encoding an inhibitor-resistant form of a plant protox enzyme, such

CC as claimed forms from wheat, soybean, cotton, sugarbeet, oilseed

CC rape, rice and sorghum (see AAW25738-48). Application of herbicide

CC will then kill undesired vegetation only. Protox enzymes can also

CC be expressed in transformed host cells and used to identify

CC inhibitors of protox enzyme activity, i.e. herbicide candidates, or

CC to design herbicide tolerant forms of the enzyme.

XX

XX Sequence 539 AA;

XX

Query Match 0.5%; Score 7; DB 18; Length 539;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112

Db 48 sskidgg 54

|||||

RESULT 187

AAW72907

ID AAW72907 standard; Protein; 539 AA.

XX

AC AAW72907;

XX

DT 16-MAY-2001 (first entry)

XX

DE Cotton protoporphyrinogen oxidase SEQ ID NO: 16.

XX

KW Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;

KW soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.

XX

OS Gossypium hirsutum.

XX

PN WO200112825-A1.

XX

XX

PD 22-FEB-2001.

XX

XX

PF 30-JUN-2000; 2000WO-EP06127.

XX

XX

PR 13-AUG-1999; 99US-0373691.

XX

XX (SYNG-) SYNGENTA PARTICIPATIONS AG.

XX

XX Johnson MA, Volrath SL, Helfetz PB, Law MD;

XX

XX WPI; 2001-234914/24.

DR

DR N-PSDB; AAF76579.

XX

XX

PT Plant DNA molecules encoding herbicide-tolerant forms of

PT protoporphyrinogen oxidase which are useful for rationally designing

PT new inhibitory herbicides and for producing herbicide-tolerant

PT transgenic plants and seeds -

XX

PS Claim 22; Page 184-187; 228pp; English.

XX The present invention provides the protein and coding sequences of a  
 CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,  
 CC sugar beet, sugar cane, cotton and oilseed rape protoporpyrinogen  
 CC oxidase (protox) enzyme. Examples of these mutants are shown in  
 CC AAB72920-AAB72926. They are useful as they enable the production of  
 CC herbicide-tolerant plants and seeds. The present sequence is a protox  
 CC protein.  
 XX  
 SQ Sequence 539 AA;

Query Match 0.5%; Score 7; DB 22; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112  
 Db 48 sskidgg 54

RESULT 188  
 AAB72921  
 ID AAB72921 standard; Protein; 539 AA.  
 XX  
 AC AAB72921;  
 XX  
 DT 16-MAY-2001 (first entry)  
 XX  
 DE Cotton protoporpyrinogen oxidase Y428H mutant.  
 XX  
 KW Protoporpyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;  
 KW soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.  
 XX  
 OS Gossypium hirsutum.  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH Misc-difference 428  
 FT /note= "wild-type Tyr substituted by His"  
 FT  
 PN WO200112825-A1.  
 XX  
 PD 22-FEB-2001.  
 XX

XX 30-JUN-2000; 2000WO-EP06127.

XX 13-AUG-1999; 99US-0373691.

XX (SYNG-) SYNGENTA PARTICIPATIONS AG.

XX Johnson MA, Volrath SL, Heifetz PB, Law MD;

XX WPI; 2001-234914/24.

XX Plant DNA molecules encoding herbicide-tolerant forms of  
 PT protoporpyrinogen oxidase which are useful for rationally designing  
 PT new inhibitory herbicides and for producing herbicide-tolerant  
 PT transgenic plants and seeds -

XX Claim 22; Page -: 228pp; English.

XX The present invention provides the protein and coding sequences of a  
 CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,  
 CC sugar beet, sugar cane, cotton and oilseed rape protoporpyrinogen  
 CC oxidase (protox) enzyme. Examples of these mutants are shown in  
 CC AAB72920-AAB72926. They are useful as they enable the production of  
 CC herbicide-tolerant plants and seeds. The present sequence is a mutant  
 CC protox protein.

CC Note: The present sequence is not shown in the specification but is  
 CC derived from that given in SEQ ID NO: 16 (see AAB72907).

XX Sequence 539 AA;

Query Match 0.5%; Score 7; DB 22; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112  
 Db 48 sskidgg 54

RESULT 189  
 AAB72922  
 ID AAB72922 standard; Protein; 539 AA.  
 XX

AC AAB72922;

XX 16-MAY-2001 (first entry)

XX Cotton protoporpyrinogen oxidase Y428A mutant.

XX Protoporpyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;  
 KW soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.

XX Gossypium hirsutum.  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH Misc-difference 428  
 FT /note= "wild-type Tyr substituted by Ala"

XX WO200112825-A1.

XX 22-FEB-2001.

XX 30-JUN-2000; 2000WO-EP06127.

XX 13-AUG-1999; 99US-0373691.

XX (SYNG-) SYNGENTA PARTICIPATIONS AG.

XX Johnson MA, Volrath SL, Heifetz PB, Law MD;

XX WPI; 2001-234914/24.

XX Plant DNA molecules encoding herbicide-tolerant forms of  
 PT protoporpyrinogen oxidase which are useful for rationally designing  
 PT new inhibitory herbicides and for producing herbicide-tolerant  
 PT transgenic plants and seeds -

XX Claim 22; Page -: 228pp; English.

XX The present invention provides the protein and coding sequences of a  
 CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,  
 CC sugar beet, sugar cane, cotton and oilseed rape protoporpyrinogen  
 CC oxidase (protox) enzyme. Examples of these mutants are shown in  
 CC AAB72920-AAB72926. They are useful as they enable the production of  
 CC herbicide-tolerant plants and seeds. The present sequence is a mutant  
 CC protox protein.

XX Note: The present sequence is not shown in the specification but is  
 CC derived from that given in SEQ ID NO: 16 (see AAB72907).

XX Sequence 539 AA;

Query Match 0.5%; Score 7; DB 22; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SSKIDGG 112  
 Db 48 sskidgg 54



## RESULT 190

AAW74581  
ID AAW74581 standard; Protein; 564 AA.

AC AAW74581;

XX DT 17-DEC-1998 (first entry)

DE DE fragment of membrane protein BA0306.

XX KW Membrane protein; BA0306; BA2303; arteriosclerosis; coronary restenosis;  
KW therapy; rabbit.

XX OS Oryctolagus sp.

FH Key Location/Qualifiers

FT Misc-difference 24 /note= "unspecified amino acid"

FT Misc-difference 146 /note= "unspecified amino acid"

FT Misc-difference 357 /note= "unspecified amino acid"

FT Misc-difference 498 /note= "unspecified amino acid"

FT /note= "unspecified amino acid"

XX WO9838305-A1.

PN 03-SEP-1998.

XX 27-FEB-1998; 98WO-JP00835.

XX 25-FEB-1998; 98JP-0062263.

PR 28-FEB-1997; 97JP-0062259.

XX (NIBS) JAPAN TOBACCO INC.

PI Nakamura Y, Tanaka T, Tsukada S;

DR WPI: 1998-481206/41.

XX N-PSDB; AAW54141.

PT Membrane protein(s) BA0306 and BA2303 - useful for, e.g. treatment  
PT and prevention of arteriosclerosis and restenosis

PS Claim 30; Page 112-115; 141pp; Japanese.

XX This sequence represents the 5' fragment of the rabbit membrane protein  
CC BA0306. DNA encoding this sequence was used to isolate the human  
CC membrane proteins BA0306 and BA2303 of the invention. The two membrane  
CC proteins are specifically expressed in mammals during arteriosclerosis  
CC and coronary restenosis. The membrane proteins, fragments of them, and  
CC antibodies against them are useful in the treatment and prevention of  
CC arteriosclerosis and restenosis. Transgenic mice expressing the  
CC extracellular region of the membrane proteins are useful as models for  
CC studying these disorders.

XX Sequence 564 AA;

Query Match 0.5%; Score 7; DB 19; Length 564;  
Best Local Similarity 100.0%; Pred. No. 5.3e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 996 AKRQAL 1002

DB 63 akrldal 69

## RESULT 191

AAW50574  
ID AAW50574 standard; Protein; 581 AA.

XX

AC AAG50574;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 64106.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 27-MAY-1999; 99US-0136021.

PR 28-MAY-1999; 99US-0136392.

PR 01-JUN-1999; 99US-0136782.

PR 03-JUN-1999; 99US-0137222.

PR 04-JUN-1999; 99US-0137528.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

```
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156559.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 581;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHLG 1174
Db 404 vsynhlg 410

RESULT 192
AAG25738
ID AAG25738 standard; Protein; 592 AA.
XX
AC AAG25738;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29926.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
```

KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0126789.

PR 06-APR-1999; 99US-0127462.

PR 08-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0129845.

PR 21-APR-1999; 99US-0130077.

PR 23-APR-1999; 99US-0130449.

PR 28-APR-1999; 99US-0130510.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132048.

PR 05-MAY-1999; 99US-0132407.

PR 06-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134236.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136382.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139452.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139750.

PR 21-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 05-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 06-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.

PR 30-AUG-1999; 99US-0151080.

PR 31-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161922.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 592;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 522 GGFNTLD 528  
 Db 526 ggfntld 532

RESULT 193

AAG37946

ID AAG37946 standard; Protein; 592 AA.

AC AAG37946;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 46742.  
 XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

OS Arabidopsis thaliana.

XX OS

XX PN

XX EP1033405-A2.

PD 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-0301439.  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123348.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 08-APR-1999; 99US-0128234.  
 PR 16-APR-1999; 99US-0128714.  
 PR 19-APR-1999; 99US-0129845.  
 PR 21-APR-1999; 99US-0130077.  
 PR 23-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 06-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 21-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 20-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144684.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 02-AUG-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147193.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151085.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 18-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160983.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161992.
PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.
Query Match 0.5%; Score 7; DB 21; Length 592;	
Best Local Similarity 100.0%; Pred. No. 5.5e+02;	
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy 522 GGFNTLD 528	
Db 526 ggfntld 532	
RESULT 194	
AAG39755	
ID AAG39755 standard; Protein; 592 AA.	
XX AC AAG39755;	
XX DT 18-OCT-2000 (first entry)	
XX Arabidopsis thaliana protein fragment SEQ ID NO: 49238.	
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49238.	
XX Protein identification; signal transduction pathway; metabolic pathway;	
KW hybridisation assay; genetic mapping; gene expression control; promoter;	
KW termination sequence.	
XX Arabidopsis thaliana.	
XX PN EF1033405-A2.	
XX PD 06-SEP-2000.	
XX 25-FEB-2000; 2000EP-0301439.	
XX 25-FEB-1999; 99US-0121825.	
PR 05-MAR-1999; 99US-0123180.	
PR 09-MAR-1999; 99US-0123548.	

PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140931.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145318.  
PR 27-JUL-1999; 99US-0145319.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 13-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0149930.  
PR 26-AUG-1999; 99US-0150566.  
PR 27-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161922.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 592;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 522 GGFNTLD 528  
 |||||  
 Db 526 ggfntld 532

RESULT 195  
 AAW74579  
 ID AAW74579 standard; Protein; 594 AA.  
 AC AAW74579;  
 DT 17-DEC-1998 (first entry)

XX Rabbit membrane protein BA2303.  
 XX Membrane protein; BA0306; BA2303; arteriosclerosis; coronary restenosis;  
 KW therapy; mouse.  
 XX Oryctolagus sp.

XX Key Location/Qualifiers  
 FT Misc-difference 111 /note= "unspecified amino acid"

XX WO9838305-A1.  
 XX 03-SEP-1998.  
 XX 27-FEB-1998; 98WO-JP00835.  
 XX 25-FEB-1998; 98JP-0062263.  
 XX 28-FEB-1997; 97JP-0062259.  
 XX (NISB ) JAPAN TOBACCO INC.  
 XX Nakamura Y, Tanaka T, Tsukada S;  
 XX PI

DR WPI: 1998-481206/41.  
 DR N-PSDB; AAV54123.  
 XX Membrane protein(s) BA0306 and BA2303 - useful for, e.g. treatment  
 PT and prevention of arteriosclerosis and restenosis  
 XX Example 4; Page 80-85; 141pp; Japanese.

XX This sequence is the rabbit membrane protein BA0306.  
 CC The DNA encoding this sequence was used to isolate the human  
 CC membrane proteins BA0306 and BA2303 of the invention. The two membrane  
 CC proteins are specifically expressed in mammals during arteriosclerosis  
 CC and coronary restenosis. The membrane proteins, fragments of them, and  
 CC antibodies against them are useful in the treatment and prevention of  
 CC arteriosclerosis and restenosis. Transgenic mice expressing the  
 CC extracellular region of the membrane proteins are useful as models for  
 CC studying these disorders.

SQ Sequence 594 AA:

Query Match 0.5%; Score 7; DB 19; Length 594;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 996 AKRLQAL 1002  
 |||||  
 Db 63 akrldal 69

RESULT 196

AAW74580  
 ID AAW74580 standard; Protein; 594 AA.  
 AC AAW74580;  
 DT 17-DEC-1998 (first entry)

XX Human membrane protein BA0306.  
 XX Membrane protein; BA0306; BA2303; arteriosclerosis; coronary restenosis;  
 KW therapy; human.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FT Misc-difference 97 /note= "unspecified amino acid"  
 FT Misc-difference 339 /note= "unspecified amino acid"  
 FT Misc-difference 498 /note= "unspecified amino acid"  
 FT Misc-difference 514 /note= "unspecified amino acid"  
 FT Misc-difference 528 /note= "unspecified amino acid"  
 FT Misc-difference 537 /note= "unspecified amino acid"  
 FT Misc-difference 565 /note= "unspecified amino acid"

XX WO9838305-A1.  
 XX 03-SEP-1998.  
 XX 27-FEB-1998; 98WO-JP00835.  
 XX 25-FEB-1998; 98JP-0062263.  
 XX 28-FEB-1997; 97JP-0062259.  
 XX (NISB ) JAPAN TOBACCO INC.  
 XX Nakamura Y, Tanaka T, Tsukada S;  
 XX PI

XX WPI; 1998-481206/41.  
 DR N-PSDB; AAV54124.  
 XX  
 PT Membrane protein(s) BA0306 and BA2303 - useful for, e.g. treatment  
 FT and prevention of arteriosclerosis and restenosis  
 XX  
 PS Claim 14; Page 90-95; 141pp; Japanese.  
 XX  
 CC This sequence is the human BA0306 membrane protein of the invention.  
 CC The invention also relates to the human BA2303 membrane protein. The two  
 CC membrane proteins are specifically expressed in mammals during  
 CC arteriosclerosis and coronary restenosis. The membrane proteins,  
 CC fragments of them, and antibodies against them are useful in the  
 CC treatment and prevention of arteriosclerosis and restenosis. Transgenic  
 CC mice expressing the extracellular region of the membrane proteins are  
 CC useful as models for studying these disorders.  
 XX  
 SQ Sequence 594 AA;  
 Query Match 0.5%; Score 7; DB 19; Length 594;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 996 AKPLQAL 1002  
 Db 63 akrqlal 69  
 |||||  
 RESULT 197  
 AAY31655  
 ID AAY31655 standard; Protein; 612 AA.  
 XX  
 AC AAY31655;  
 XX  
 DT 09-NOV-1999 (first entry)  
 DE  
 DE HLA-DR2 alpha-Fos-IgM fusion protein.  
 XX  
 KW Major histocompatibility complex Class II; MHC; binding domain;  
 KW HLA-DR2; leucine zipper; Fos; IgM; Fc; immunoglobulin; antibody;  
 KW fusion protein; multiple sclerosis; rheumatoid arthritis;  
 KW graft rejection; allergy; autoimmune disease; pemphigus vulgaris;  
 KW systemic lupus erythematosus; T lymphocyte; T cell; diagnosis;  
 KW therapy; adoptive immunotherapy.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Saccharomyces cerevisiae.  
 OS Chimeric - synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..25  
 FT /note= "alpha-mating factor secretion signal"  
 FT Protein 26..612  
 FT /note= "mature protein"  
 FT Domain 26..217  
 FT /note= "DRA\*0101 extracellular domain"  
 FT Peptide 218..223  
 FT /note= "linker"  
 FT Domain 224..263  
 FT /note= "Fos leucine zipper domain"  
 FT Domain 264..612  
 FT /note= "IgM"  
 XX  
 PN WO9942597-A1.  
 XX  
 PD 26-AUG-1999.  
 XX  
 PF 19-FEB-1999; 99WO-US03603.  
 XX  
 PR 19-FEB-1998; 98US-0075351.  
 XX

PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Strominger JL, Wucherpfennig KW;  
 XX  
 DR WPI; 1995-527481/44.  
 DR N-PSDB; AAX87814.  
 XX  
 PT New MHC Class II binding domain fusion proteins and conjugates -  
 FT used for, e.g. treating allergic and autoimmune diseases or  
 FT detecting, isolating, activating or killing specific T cells  
 XX  
 PS Example 7; Page 105-107; 113pp; English.  
 XX  
 CC The present sequence represents a decavalent HLA-DR2 MHC binding  
 CC domain fusion protein comprising an alpha-mating factor secretion  
 CC signal, the extracellular domain of the HLA-DR2 alpha chain  
 CC (residues 1-191 of DRA\*0101), a 7-amino acid linker, the 40-amino  
 CC acid leucine zipper dimerization domain of Fos, and the Fc portion  
 CC of IgM. The invention provides new monovalent, multivalent and  
 CC multimeric MHC Class II binding domain fusion proteins and  
 CC conjugates comprising at least a binding domain of an MHC Class II  
 CC alpha or beta chain and a dimerization domain. The MHC fusion  
 CC proteins and conjugates can be used: for detecting and isolating T  
 CC cells having a defined MHC/peptide complex specificity (claimed);  
 CC to confer to a subject adoptive immunity to a defined MHC/peptide  
 CC complex (claimed); to stimulate or activate T cells reactive to a  
 CC defined MHC/peptide complex (claimed); for selective killing of T  
 CC cells reactive to a defined MHC complex (claimed); to tolerate a  
 CC subject to a defined MHC/peptide complex (claimed); to treat  
 CC allergic and autoimmune diseases, e.g. multiple sclerosis,  
 CC rheumatoid arthritis, pemphigus vulgaris, and systemic lupus  
 CC erythematosus; and to prevent organ or tissue transplant rejection.  
 CC The DR2-IgM design may be particularly useful for immunotherapy  
 CC because of: higher affinity for the T cell receptors on cognate  
 CC T cells; complement fixation by the Fc segment of IgM; and longer  
 CC serum half-life.  
 XX  
 SQ Sequence 612 AA;  
 Query Match 0.5%; Score 7; DB 20; Length 612;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1109 FEAQGA 1115  
 Db 79 feaqga 85  
 |||||  
 RESULT 198  
 AAY81659  
 ID AAY81659 standard; Protein; 628 AA.  
 XX  
 AC AAY81659;  
 XX  
 DT 24-MAY-2000 (first entry)  
 DE  
 DE Streptococcus pneumoniae protein sequence ID308.  
 XX  
 KW Streptococcus pneumoniae; vaccine; screening; protein antigen;  
 KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;  
 KW pneumococcal disease.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN WO200006737-A2.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PF 27-JUL-1999; 99WO-GB02451.  
 XX  
 PR 27-JUL-1998; 98GB-0016337.  
 PR 19-MAR-1999; 99US-0125164.  
 PR



XX (MICR-) MICROBIAL TECHNICS LTD.  
 XX PA  
 XX PI Gilbert CFG, Hansbro PM;  
 XX DR WPI; 2000-195300/17.  
 XX PT New Streptococcal protein, useful as a vaccine, for diagnosis of  
 PT pneumococcal diseases and for screening agents capable of antagonizing  
 PT or inhibiting expression of the protein  
 XX PS Claim 2; Page 98; 108pp; English.  
 XX CC AAY81501 to AAY81679 represent specifically claimed protein sequences  
 CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent  
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.  
 CC The sequences have antibacterial and antiinflammatory properties.  
 CC The protein sequences, and fragments of them, are useful as immunogens  
 CC and/or antigens. The nucleotide sequences can be used in vaccines and in  
 CC diagnostic assays. The proteins and nucleotides can be useful for the  
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also  
 CC useful for screening an agent capable of antagonising, inhibiting or  
 CC interfering with the function or expression of the proteins in which the  
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection  
 CC and meningitis. AAA05591 to AAA05614 represent primers used in the  
 CC exemplification of the present invention.  
 XX SQ Sequence 628 AA;

Query Match 0.5%; Score 7; DB 21; Length 628;  
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SLALVGA 22  
 |||||  
 Db 7 slalvga 13

RESULT 199  
 AAG50573  
 ID AAG50573 standard; Protein; 645 AA.  
 AC AAG50573;  
 XX  
 XX 18-OCT-2000 (first entry)  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 64105.  
 DE Arabidopsis thaliana.  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 OS Arabidopsis thaliana.  
 PN EP1033405-A2.  
 XX  
 XX 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.

```
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157553.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 7; DB 21; Length 645;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 VSYNHlg 1174
Db 468 vsynhlg 474

RESULT 200
AAG25737
ID AAG25737 standard; Protein; 673 AA.
XX
AC AAG25737;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29925.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
```

PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145216.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157753.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match . . . . . 0.5%; Score 7; DB 21; Length 673;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 522 GGFNTLD 528

Db 607 gsfntld 613

Search completed: August 29, 2001, 09:35:15  
Job time: 126 sec

FT PROPEP ? 1310 POTENTIAL.  
 SQ SEQUENCE 1310 AA; 141988 MW; 1BC21FE3D435F981 CRC64;  
  
 Query Match 4.8%; Score 62; DB 1; Length 1310;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-55;  
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1082 ANSLNSGANTNFVYTSRIIFANQHEHFEAQQGALGSDQSSLNFKSALLQDLNOSYHYLAY 1141  
 |||||  
 Db 1097 ANSLNSGANTNFVYTSRIIFANQHEHFEAQQGALGSDQSSLNFKSALLQDLNOSYHYLAY 1156  
 |||||  
 QY 1142 SA 1143  
 ||  
 Db 1157 SA 1158  
  
 RESULT 7  
 ID ATPX CYAPA STANDARD; PRT; 164 AA.  
 AC P48085;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE ATP SYNTHASE B' CHAIN (EC 3.6.1.34) (SUBUNIT II).  
 ATPG.  
 GN Cyanophora paradoxa.  
 OS  
 OG Cyanelle.  
 OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.  
 NCBI\_TaxID=2762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LB555 / PRINGSHEIM;  
 RA Stirewalt V.L., Michalowski C.B., Luffelhardt W., Bohnert H.J.,  
 Bryant D.A.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: THIS IS ONE OF THE FOUR CHAINS OF THE NONENZYMATIC  
 COMPONENT (CF(0) SUBUNIT) OF THE CHLOROPLAST ATPASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: CYANELLE THYLAKOID MEMBRANE.  
 CC -1- SIMILARITY: THE B' SUBUNIT IS A DIVERGED AND DUPLICATED FORM OF  
 B FOUND IN PLANTS AND PHOTOSYNTHETIC BACTERIA.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; U30821; AAA81256.1; -  
 Dr Mendel; 7860; CYAPA; atpg;1.  
 Dr InterPro; IPR002146; -  
 Dr Pfam; PF00430; ATP-synt\_B; 1.  
 KW Hydrogen ion transport; Transmembrane; CF(0); Cyanelle.  
 SQ SEQUENCE 164 AA; 18568 MW; 49AAACE15AF010D7C CRC64;  
  
 Query Match 0.6%; Score 8; DB 1; Length 164;  
 Best Local Similarity 100.0%; Pred. No. 3.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 715 FYRPLIKI 722  
 |||||  
 Db 46 FYRPLIKI 53  
  
 RESULT 8  
 ID VATE-SULSO STANDARD; PRT; 194 AA.  
 AC Q9UWW5;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)

01-OCT-2000 (Rel. 40, Last annotation update)  
 V-TYPE ATP SYNTHASE SUBUNIT E (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT E).  
 ATPE.  
 Sulfolobus solfataricus.  
 Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
 NCBI\_TaxID=2287;  
 [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN-DSM 1617 / P2;  
 RX MEDLINE=20165948; PubMed=10701121;  
 RA Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,  
 RA Confalonieri F., Curtis B., Duguet M., Erosa G., Paquy D.,  
 RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,  
 RA Kishwana N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,  
 RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,  
 RA Ragan M.A., Sengen C.W.;  
 RT "Gene content and organization of a 281-kbp contig from the genome of  
 RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";  
 RL Genome 43:116-136(2000).  
 CC !- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
 CC !- GRADIENT ACROSS THE MEMBRANE.  
 CC !- SIMILARITY: BELONGS TO THE V-ATPASE E SUBUNIT FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Y18930; CAB57738.1;  
 DR Hydrolyase; ATP synthetase; Hydrogen ion transport.  
 KW SEQUENCE 194 AA; 22633 MW; 057157C921ED8449 CRC64;  
 SQ  
 Query Match 0.6%; Score 8; DB 1; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 555 INELVKT 562  
 Db 48 INELVKT 55  
 RESULT 9  
 ARGCBUCAI STANDARD; PRT; 334 AA.  
 AC 57156;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (EC 1.2.1.38) (AGPR) (N-  
 DE ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE).  
 GN ARGCB OR BU048.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TOKYO 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 RT Buchnera sp. APS.";  
 RL Nature 407:81-86(2000).  
 CC !- CATALYTIC ACTIVITY: N-ACETYL-L-GLUTAMATE 5-SEMIALDEHYDE + NADP(+) +  
 CC + ORTHOPHOSPHATE -> N-ACETYL-5-GLUTAMYL PHOSPHATE + NADPH.  
 CC !- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.  
 CC !- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

Aug 29 10:20:38 2001

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AP001118; BAB12771.1;  
 DR PROSITE; PS01224; ARGCC; 1;  
 KW Arginine biosynthesis; Oxidoreductase; NADP.  
 FT ACT\_SITE 154 154 BY SIMILARITY.  
 SQ SEQUENCE 334 AA; 37815 MW; 1032E60048AFA90A CRC64;  
 Query Match 0.6%; Score 8; DB 1; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 209 SGAGRKAS 216  
 Db 187 SGAGRKAS 194  
 RESULT 10  
 ARGCBACSU STANDARD; PRT; 346 AA.  
 AC P23715; P70953; O08146;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (EC 1.2.1.38) (AGPR) (N-  
 DE ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE).  
 GN ARGCC  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168 / EMG50;  
 RX MEDLINE=90356403; PubMed=2117746;  
 RA Smith M.C.M., Mountain A., Baumberg S.;  
 RT "Nucleotide sequence of the Bacillus subtilis argC gene encoding N-  
 RT acetylglutamate-gamma-semialdehyde dehydrogenase.";  
 RL Nucleic Acids Res. 18:4595-4595(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE=94297722; PubMed=8025667;  
 RA O'Reilly M., Devine K.M.;  
 RT "Sequence and analysis of the citrulline biosynthetic operon argC-F  
 RT from Bacillus subtilis.";  
 RL Microbiology 140:1023-1025(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE=97177785; PubMed=9025291;  
 RA Levine A., Vannier F., Roche B., Autret S., Mavel D., Seror S.J.;  
 RT "A 10.3 kbp segment from nprB to argJ at the 102 degrees region of  
 RT the Bacillus subtilis chromosome.";  
 RL Microbiology 143:175-177(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE=98015415; PubMed=9353931;  
 RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;  
 RT "Sequencing of regions downstream of addA (98 degrees) and citG (289  
 RT degrees) in Bacillus subtilis.";  
 RL Microbiology 143:3305-3308(1997).  
 RN [5]  
 RP SEQUENCE OF 1-56 FROM N.A.  
 RX MEDLINE=87192000; PubMed=3106155;  
 RA Smith M.C.M., Mountain A., Baumberg S.;

BL + C  
- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.  
- SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

\*Sequence analysis of the Bacillus subtilis argC promoter region.\*;  
RL Gene 49:53-60(1986).  
CC -1- CATALYTIC ACTIVITY: N-ACETYL-L-GLUTAMATE 5-SEMIALDEHYDE + NADP(+) + ORTHOPHOSPHATE -> N-ACETYL-5-GLUTAMYL PHOSPHATE + NADPH.  
CC -1- PATHWAY: THIRD STEP IN ARGinine BIOSYNTHESIS.  
CC -1- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X52834; CAA37016.1;  
DR EMBL; 226919; CAA81543.1;  
DR EMBL; 279580; CAB01842.1;  
DR EMBL; Y09476; CAA70638.1;  
DR EMBL; Z99110; CAB12976.1;  
DR EMBL; Z99109; CAB12960.1;  
DR EMBL; M15420; AAA22248.1;  
DR PIR; S12592; S12592;  
DR Subtilisin; BG10191; argC.  
DR InterPro; IPR000534;  
DR InterPro; IPR000706;  
DR Pfam; PF01118; Semialdehyde\_dh; 1.  
DR PROSITE; PS01224; ARG; 1.  
KW Arginine biosynthesis; Oxidoreductase; NADP.  
FT ACT\_SITE 149 149 BY SIMILARITY.  
FT CONFLICT 235 235 F -> V (IN REF. 3 AND 4).  
FT CONFLICT 341 341 MISSING (IN REF. 1).  
SQ SEQUENCE 346 AA; 38121 MW; 9E22F2AB31B7542B CRC64;

Query Match 0.6%; Score 8; DB 1; Length 346;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKAS 216  
DB 181 SGAGRKAS 188  
|||||

RESULT 11  
DMS NEIMB STANDARD; PRT; 637 AA.

AC Q9JW13;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE 1-DEOXY-D-XYLULOSE 5-PHOSPHATE SYNTHASE (EC 2.2.-.-) (1-DEOXYXYLULOSE-5-PHOSPHATE SYNTHASE) (DXP SYNTHASE) (DXPS).  
GN DMS OR NMA0589.  
OS Neisseria meningitidis (serogroup A).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=65699;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22491 / SEROGROUP A / SEROTYPE 4A;  
RX MEDLINE=2022556; PubMed=10761919;  
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C., Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T., Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S., Jagers K., Leather S., Moule S., Mungall K., Quail M.A., Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J., Whitehead S., Spratt B.G., Barrall B.G.;  
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491";  
RL Nature 404:502-506(2000).  
CC -1- FUNCTION: CATALYZES THE ACYLATION CONDENSATION REACTION BETWEEN C ATOMS 2 AND 3 OF PYRUVATE AND GLYCERALDEHYDE 3-PHOSPHATE TO YIELD 1-DEOXY-D-XYLULOSE-5-PHOSPHATE (DXP) (BY SIMILARITY).

CC -1- COFACTOR: THIAMINE PYRROPHOSPHATE (BY SIMILARITY).  
CC -1- PATHWAY: DEOXYXYLULOSE-5-PHOSPHATE PATHWAY (DXP) OF ISOPRENOID BIOSYNTHESIS; FIRST STEP. BIOSYNTHETIC PATHWAY TO THIAMINE AND PYRIDOXOL; FIRST STEP.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY. DMS SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AL162753; CAB83880.1;  
DR InterPro; IPR000360;  
DR PROSITE; PS00801; TRANSKETOLASE\_1; 1.  
DR PROSITE; PS00802; TRANSKETOLASE\_2; 1.  
KW Transferase; Flavoprotein; Thiamine pyrophosphate;  
KW Isoprene biosynthesis; Thiamine biosynthesis.  
SQ SEQUENCE 637 AA; 68720 MW; 3B2BBD01AAD182F5 CRC64;

Query Match 0.6%; Score 8; DB 1; Length 637;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 GHFASNLG 1291  
DB 44 GHFASNLG 51  
|||||

RESULT 12  
DMS NEIMB STANDARD; PRT; 637 AA.

AC Q9JXV7;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE 1-DEOXY-D-XYLULOSE 5-PHOSPHATE SYNTHASE (EC 2.2.-.-) (1-DEOXYXYLULOSE-5-PHOSPHATE SYNTHASE) (DXP SYNTHASE) (DXPS).  
GN DMS OR NMB1867.  
OS Neisseria meningitidis (serogroup B).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=WC58 / SEROGROUP B;  
RX MEDLINE=20175755; PubMed=10710307;  
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J., Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;  
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58";  
RL Science 287:1809-1815(2000).  
CC -1- FUNCTION: CATALYZES THE ACYLATION CONDENSATION REACTION BETWEEN C ATOMS 2 AND 3 OF PYRUVATE AND GLYCERALDEHYDE 3-PHOSPHATE TO YIELD 1-DEOXY-D-XYLULOSE-5-PHOSPHATE (DXP) (BY SIMILARITY).  
CC -1- COFACTOR: THIAMINE PYRROPHOSPHATE (BY SIMILARITY).  
CC -1- PATHWAY: DEOXYXYLULOSE-5-PHOSPHATE PATHWAY (DXP) OF ISOPRENOID BIOSYNTHESIS; FIRST STEP. BIOSYNTHETIC PATHWAY TO THIAMINE AND PYRIDOXOL; FIRST STEP.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY. DMS SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -





PIR 68

Title: Divergence of genetic sequences for the vacuolating cytotoxin among Helicobacter  
Reference number: A53739; MUID:94193753  
A:Accession: B53739  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1287 <COV>  
A:Cross-references: GB:005676; NID:9471727; PIDN:AAAL7657.1; PID:9471729  
A:Note: parts of this sequence, including the amino end of the mature protein, were determined by J. Biol. Chem. 267, 10570-10575, 1992  
A:Title: Purification and characterization of the vacuolating toxin from Helicobacter pylori  
A:Reference number: A38137; MUID:92268100  
A:Accession: A38137  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 34-56 <CO2>  
A:Note: sequence extracted from NCBI backbone (NCBIP:103729)  
C:Genetics:  
A:Gene: vacA  
C:Words: cytotoxin  
F:1-1290/Product: signal sequence #status predicted <SIG>

Query Match 7.6%; Score 99; DB 2; Length 1287;  
Best Local Similarity 100.0%; Pred. No. 1e-94;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 697 VGNAAAMFNNDIDSATGFKPLIKINSADLIKNTKTHVAKKIIIGYGVNSTGNGISN 756  
Db 689 VGNAAAMFNNDIDSATGFKPLIKINSADLIKNTKTHVAKKIIIGYGVNSTGNGISN 748  
QY 757 VNLSEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 795  
Db 749 VNLSEQFKERLALYNNNNRMDTCVVRNTDDIKACGMAIG 787

RESULT 3  
G64630  
vacuolating cytotoxin precursor - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: G64630  
R:Tomb, J.F.; White, O.; Kervage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: G64630  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1290 <TON>  
A:Cross-references: GB:AE000598; GB:AE000511; NID:92314019; PIDN:AA007935.1; PID:9231402  
F:1-1290/Product: signal sequence #status predicted <SIG>  
F:34-1290/Product: vacuolating cytotoxin #status predicted <MAT>

Query Match 6.7%; Score 87; DB 2; Length 1290;  
Best Local Similarity 100.0%; Pred. No. 4.3e-82;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 447 VENLTGNTVDGRLRVNNOVGVALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 506  
Db 442 VENLTGNTVDGRLRVNNOVGVALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 501  
QY 507 VDAHTANFKGIDTNGGFGNTLDFSGVT 533  
Db 502 VDAHTANFKGIDTNGGFGNTLDFSGVT 528

RESULT 4  
S44983  
A:Status: preliminary; translated from GB/EMBL/DBDJB  
A:Molecule type: DNA

vacuolating cytotoxin precursor - Helicobacter pylori (isolate 185-44)  
C:Species: Helicobacter pylori  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
C:Accession: S44983; S44102  
R:Schmitt, W.; Haas, R.  
Mol. Microbiol. 12, 307-319, 1994  
A:Title: Genetic analysis of the Helicobacter pylori vacuolating cytotoxin: structure and function of the mature protein  
A:Reference number: S44983; MUID:94335650  
A:Accession: S44983  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1291 <SCH>  
A:Cross-references: EMBL:Z26883; NID:9472941; PIDN:CAA81528.1; PID:9472942  
C:Keywords: cytotoxin  
F:1-1291/Product: signal sequence #status predicted <SIG>  
F:34-1291/Product: vacuolating cytotoxin #status predicted <MAT>

Query Match 6.7%; Score 87; DB 2; Length 1291;  
Best Local Similarity 100.0%; Pred. No. 4.3e-82;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 447 VENLTGNTVDGRLRVNNOVGVALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 506  
Db 442 VENLTGNTVDGRLRVNNOVGVALAGSSANFEKAGTDTKNGTATFNNDISLGRFVNLK 501  
QY 507 VDAHTANFKGIDTNGGFGNTLDFSGVT 533  
Db 502 VDAHTANFKGIDTNGGFGNTLDFSGVT 528

RESULT 5  
A53739  
hypothetical vacuolating cytotoxin - Helicobacter pylori (strain ATCC 53726) (fragment)  
C:Species: Helicobacter pylori  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
C:Accession: A53739  
R:Cover, T.L.; Tummuru, M.K.R.; Cao, P.; Thompson, S.A.; Blaser, M.J.  
J. Biol. Chem. 269, 10566-10577, 1994  
A:Title: Divergence of genetic sequences for the vacuolating cytotoxin among Helicobacter pylori strains  
A:Reference number: A53739; MUID:94193753  
A:Accession: A53739  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-513 <COV>  
A:Cross-references: GB:J05677; NID:9471730; PIDN:AAAL7658.1; PID:9471731  
A:Note: This strain, designated 87-203, ATCC 53726 tox-, does not possess toxin activity  
C:Genetics:  
A:Gene: vacA

Query Match 1.5%; Score 20; DB 2; Length 513;  
Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 381 QPTQV DGPFGAGGKDTVNI 400  
Db 44 QPTQV DGPFGAGGKDTVNI 63

RESULT 6  
T06913  
H+transferring ATP synthase (EC 3.6.1.34) chain b' - Cyanophora paradoxa cyanelle  
C:Species: cyanelle Cyanophora paradoxa  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
C:Accession: T06913  
R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.  
submitted to the EMBL Data Library, July 1995  
A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.  
A:Reference number: Z15840  
A:Accession: T06913  
A:Status: preliminary; translated from GB/EMBL/DBDJB  
A:Molecule type: DNA

A:Residues: 1-164 <STI>  
 A:Cross-references: EMBL:U30821; NID:g1016083; PIDN:AAA81256.1; PID:g1016169  
 A:Experimental source: strain Pringsheim LB555

C:Genetics:

A:Gene: atpG

A:Genome: cyanelle

C:Superfamily: H-transferring ATP synthase chain I

C:Keywords: ATP biosynthesis; cyanelle; hydrolase; thylakoid  
 Query Match 0.6%; Score 8; DB 2; Length 164;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 715 FYKPLIKI 722

|||||

DB 46 FYKPLIKI 53

RESULT 7

C8: N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) [imported] - Buchnera sp. (str)  
 C:Species: Buchnera sp.

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: C84935

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A

A:Reference number: A84930; MUID:20445173

A:Accession: C84935

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144

A:Experimental source: strain APS

C:Genetics:

A:Gene: argC; BU048

C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase

C:Keywords: oxidoreductase

Query Match 0.6%; Score 8; DB 2; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKAS 216

|||||

DB 187 SGAGRKAS 194

RESULT 8

N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) - Bacillus subtilis  
 N:Alternate names: acetylglutamate semialdehyde dehydrogenase; N-acetylglutamate-gamma-S

C:Species: Bacillus subtilis

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000

C:Accession: I40372; S12592; A26526; S38428; F69588; S20023

R:O'Reilly, M.; Devine, K.M.

Microbiology 140, 1023-1025, 1994

A:Title: Sequence and analysis of the citrulline biosynthetic operon argC-F from Bacillu

A:Reference number: I40372; MUID:94297722

A:Accession: I40372

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-346 <RES>

A:Cross-references: EMBL:Z26919; NID:g408113; PID:g575620

R:Smith, M.C.M.; Mountain, A.; Baumberg, S.

Nucleic Acids Res. 18, 4595, 1990

A:Title: Nucleotide sequence of the Bacillus subtilis argC gene encoding N-acetylglutamate

A:Reference number: S12592; MUID:90356403

A:Accession: S12592

A:Molecule type: DNA

A:Residues: 1-340,342-346 <SMI>

A:Cross-references: EMBL:X52834; NID:g39806; PIDN:CAA37016.1; PID:g580828  
 R:Smith, M.C.M.; Mountain, A.; Baumberg, S.  
 Gene 49, 53-60, 1986  
 A:Title: Sequence analysis of the Bacillus subtilis argC promoter region.  
 A:Reference number: A26526; MUID:87192000  
 A:Accession: A26526  
 A:Molecule type: DNA  
 A:Residues: 1-56 <SM2>  
 A:Cross-references: GB:M15420; NID:g142533; PIDN:AAA22248.1; PID:g142534  
 R:O'Reilly, M.; Devine, K.M.  
 submitted to the EMBL Data Library, October 1993  
 A:Reference number: S38428  
 A:Accession: S38428  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-346 <OAR>  
 A:Cross-references: EMBL:Z26919  
 R:Czaplewski, L.G.; North, A.K.; Smith, M.C.M.; Baumberg, S.; Stockley, P.G.  
 Mol. Microbiol. 6, 267-275, 1992  
 A:Title: Purification and initial characterization of AhrC: the regulator of arginine

A:Reference number: S20023; MUID:92186717

A:Contents: annotation

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
 C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
 A.: Ehrlich, S.D.; Emerson, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
 Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M  
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtili

A:Title: The complete genome sequence of the gram-positive bacterium

A:Reference number: A69580; MUID:98044033

A:Accession: F69588

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-234, 'V', 236-346 <KUN>

A:Cross-references: GB:Z99109; GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB12976.1;

A:Experimental source: strain 168

C:Genetics:

A:Gene: argC

A:Start codon: TTG

C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase

C:Keywords: oxidoreductase

Query Match 0.6%; Score 8; DB 1; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SGAGRKAS 216

|||||

DB 181 SGAGRKAS 188

RESULT 9

S61147

TCMU protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein D9476.9; protein YDR350c

C:Species: Saccharomyces cerevisiae

C>Date: 23-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 23-Mar-2001

C:Accession: S61147; S59675

R:Du, Z.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of S. cerevisiae cosmid 9476.

A:Reference number: S61148

A:Accession: S61147

A:Molecule type: DNA

A;Residues: 1-611 <DUZ>  
A;Cross-references: EMBL:U28372; NID:9849170; PID:9849179; MIPS:YDR350C  
R;Zhang, Y.; Robinson, K.M.; Lemire, B.D.  
submitted to the EMBL Data Library, July 1995  
A;Reference number: S59675  
A;Accession: S59675  
A;Molecule type: DNA  
A;Residues: 1-591; 'GARSWYK', 598, 'LFGGFEIRHMAIQLIIRKQGWPFKPFNFDITLLTELVENNIKEPTDSTLF', 7  
A;Cross-references: EMBL:U32306; NID:9929984; PID:9929985  
A;Experimental source: strain MH125  
C;Genetics:  
A;Gene: 'SGD:TCM10'  
A;Cross-references: SGD:S0002758; MIPS:YDR350C  
A;Map position: 4R

Query Match 0.6%; Score 8; DB 2; Length 611;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 30 LYRSLSS 107  
Db 452 LYRSLSS 459

RESULT 10  
R81978  
probable 1-deoxyxylulose-5-phosphate synthase NMA0589 [imported] - Neisseria meningitidis  
C;Species: Neisseria meningitidis  
C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C;Accession: B81978  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.  
A;Reference number: A81775; MUID:20222556  
A;Accession: B81978  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-637 <PAR>  
A;Cross-references: GB:AL162753; GB:AL157959; NID:97379120; PIDN:CAB83880.1; PID:g737932  
A;Experimental source: serogroup A, strain 22491  
C;Genetics:  
A;Gene: dxs; NMA0589  
C;Superfamily: hypothetical protein C2814

Query Match 0.6%; Score 8; DB 2; Length 637;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1284 GHFASNLG 1291  
Db 44 GHFASNLG 51

RESULT 11  
D81034  
1-deoxyxylulose-5-phosphate synthase NMB1867 [imported] - Neisseria meningitidis (strain  
C;Species: Neisseria meningitidis  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C;Accession: D81034  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzia, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: A81000; MUID:20115755  
A;Accession: D81034  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-637 <TET>

A;Cross-references: GB:AE002536; GB:AE002098; NID:97227115; PIDN:AAF42201.1; PID:g722  
A;Experimental source: serogroup B, strain MC58  
C;Genetics:  
A;Gene: NMB1867  
C;Superfamily: hypothetical protein C2814

Query Match 0.6%; Score 8; DB 2; Length 637;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 GHFASNLG 1291  
Db 44 GHFASNLG 51  
RESULT 12  
T19011  
hypothetical protein C06C6.7 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T19011  
R;McMurray, A.  
submitted to the EMBL Data Library, March 1997  
A;Reference number: Z19059  
A;Accession: T19011  
A;Status: preliminary; translated from GB/EMBL/DDBB  
A;Molecule type: DNA  
A;Residues: 1-935 <WIL>  
A;Cross-references: EMBL:Z93374; PIDN:CAB07557.1; GSPDB:GN00023; CESP:C06C6.7  
A;Experimental source: clone C06C6  
C;Genetics:  
A;Gene: CESP:C06C6.7  
A;Map position: 5  
A;Introns: 28/1; 55/1; 80/1; 801/2; 865/2

Query Match 0.6%; Score 8; DB 2; Length 935;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 20 VGLVSVIT 27  
Db 609 VGLVSVIT 616  
RESULT 13  
B70729  
hypothetical protein RV2566 - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C;Accession: B70729  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A;Reference number: A70500; MUID:98295987  
A;Accession: B70729  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1140 <COL>  
A;Cross-references: GB:Z77250; GB:AL123456; NID:g3261617; PIDN:CAB01049.1; PID:e25533  
A;Experimental source: strain H37RV  
C;Genetics:  
A;Gene: RV2566

Query Match 0.6%; Score 8; DB 2; Length 1140;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 ITVDGDLR 461  
 DB 1046 ITVDGDLR 1053

RESULT 14  
 A64556  
 toxin-like outer membrane protein HP0289 - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 C:Accession: A64556  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khairat, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: A64556  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2893 <TON>  
 A:Cross-references: GB:AE000547; GB:AE000511; MID:g2313377; PIDN:AAD07355.1; PID:g231338

Query Match 0.6%; Score 8; DB 2; Length 2893;  
 Best Local Similarity 100.0%; Pred. No. 99;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
 DB 176 NRVGSGAG 183

RESULT 15  
 C71953  
 toxin-like outer membrane protein jhp0274 - Helicobacter pylori (strain J99)  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99  
 C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
 C:Accession: C71953  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.  
 A:Reference number: A71800; MUID:99120557  
 A:Accession: C71953  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2902 <ARN>  
 A:Cross-references: GB:AF001464; GB:AE001439; MID:g4154789; PIDN:AAD05855.1; PID:g415479  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: jhp0274

Query Match 0.6%; Score 8; DB 2; Length 2902;  
 Best Local Similarity 100.0%; Pred. No. 99;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
 DB 185 NRVGSGAG 192

RESULT 16  
 D69681  
 peptidase synthetase ppsd - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Nov-2000  
 C:Accession: D69681  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

Wed Aug 29 10:20:36 2001

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galiech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seakeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
 A:Reference number: A69580; MUID:98044033  
 A:Accession: D69681  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-3603 <KUN>  
 A:Cross-references: GB:Z99113; GB:AL009126; MID:g2634090; PIDN:CABL3714.1; PID:g26342  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: ppsd  
 C:Superfamily: surfactin synthetase; acetate-CoA ligase homology; acyl carrier prote  
 C:Keywords: carrier protein; phosphopantetheine; phosphoprotein  
 F:509-952/Domain: acetate-CoA ligase homology <ACLI>  
 F:969-1037/Domain: acyl carrier protein homology <ACP1>  
 F:1540-1983/Domain: acetate-CoA ligase homology <ACLI2>  
 F:2000-2068/Domain: acyl carrier protein homology <ACP2>  
 F:2579-3019/Domain: acetate-CoA ligase homology <ACLI3>  
 F:3037-3104/Domain: acyl carrier protein homology <ACP3>  
 F:1001,2032,3069/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 0.6%; Score 8; DB 1; Length 3603;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 LTLQASEG 227  
 DB 2455 LTLQASEG 2462

## RESULT 17

PH1678

Ig heavy chain V region (clone NP-6-9) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999

C:Accession: PH1678

R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.

A:Reference number: PH1675; MUID:93301607

A:Accession: PH1678

A:Molecule type: mRNA

A:Residues: 1-22 <MCH>

A:Experimental source: B cell

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

## Query Match

Best Local Similarity 100.0%; Pred. No. 12;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1205 EARYYYG 1211

DB 12 EARYYYG 18

## RESULT 18

PH1679

Ig heavy chain V region (clone NP-6-10) - mouse (fragment)

C:Species: Mus musculus (house mouse)

infection or to identify H. pylori polypeptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors. The genomic sequence of H. pylori (ATCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.

Sequence 246 AA;

Query Match 0.6%; Score 8; DB 18; Length 246;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

205 NRVGSGAG 212  
|||||||  
185 nrvgsgag 192

RESULT 31  
AAV11003  
ID AAV11003 standard; Protein; 246 AA.

AC AAV11003;  
XX  
DT 08-JUN-1999 (first entry)  
DE H. pylori ORF hp2el10911\_960952\_c2\_86 cellular protein.  
KW Vaccine; probe; diagnostic; ORF; cell envelope protein;  
KW secreted protein; cellular protein.

OS Helicobacter pylori.

PN W09818323-Al.  
XX  
PD 07-MAY-1998.  
XX  
PF 28-OCT-1997; 97WO-US19575.  
PR 14-JUL-1997; 97US-0891928.  
PR 28-OCT-1996; 96US-0739150.  
PR 06-DEC-1996; 96US-0759739.

PA (ASTR) ASTRA AB.

XX Alm RA, Smith D;

XX WPI; 1998-271811/24.

DR N-PSDB; AAX30470.

PT Helicobacter pylori nucleic acids and proteins - used to develop products for the detection, prevention and treatment of H. pylori infections

PS Claims 27, 31; Page 213-214; 279pp; English.

XX Recombinant or substantially pure preparations of H. pylori polypeptides are disclosed, together with the nucleic acids encoding them. In all, 73 ORFs are shown. The proteins are variously cell envelope proteins, secreted proteins or other cellular proteins. Vaccines containing at least nucleic acids of proteins are claimed, as are probes containing at least 8 nucleotides from the nucleic acid sequences. The vaccines are useful for treating or reducing the risk of H. pylori infections, and the probes can be used diagnostically for detecting the presence of Helicobacter in a sample. The products are also of use in screening for compounds having the ability to interfere with the H. pylori life

CC cycle or to inhibit H. pylori infection.  
XX  
SQ Sequence 246 AA;

Query Match 0.6%; Score 8; DB 19; Length 246;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
|||||||  
DB 185 nrvgsgag 192

RESULT 32  
AAW80647  
ID AAW80647 standard; Protein; 293 AA.

XX  
AC AAW80647;  
XX  
DT 24-DEC-1998 (first entry)  
DE S. pneumoniae transport protein.

KW Streptococcus pneumoniae protein; recombinant; gene expression; DNA chip;  
KW virulence; antibody; infection; detection; treatment; hypothetical;  
KW cell wall biosynthetic, external target; minimal gene set protein.

OS Streptococcus pneumoniae.

PN W09826072-Al.

PD 18-JUN-1998.

PF 09-DEC-1997; 97WO-US22578.

PR 13-DEC-1996; 96US-0036281.

XX (ELIL) LILLY & CO ELI.

XX Baltz RH, Burgett SG, Dehoff BS, Hoskins JA, Jaskunas SR;  
PI Mills BJ, Norris FH, Peery RB, Rostock PK, Rostock PR;  
PI Skatrud PL, Smith MC, Solenberg PJ, Treadway PJ;  
PI Young Bellido ML;

XX WPI; 1998-348529/30.

DR N-PSDB; AAV65221.

PT Streptococcus pneumoniae nucleic acid sequences - used in DNA chips for evaluating gene expression, and identification of virulence genes

PS Claim 3; Pages 234-235; 333pp; English.

XX This sequence represents a Streptococcus pneumoniae transport protein. The invention provides DNA sequences (AAV65201 to AAV65304) from the Streptococcus pneumoniae genome and corresponding protein sequences (AAW80605 to AAW80728). The protein sequences are classified as hypothetical, cell wall biosynthetic, external target, or minimal gene set proteins. A recombinant host containing a vector comprising any of the above nucleic acids can be used for the recombinant expression of the proteins. The invention also provides a DNA chip having arrayed on it at least 15 base pair fragment of any one or more of these DNA sequences. The DNA chip can be used for evaluating gene expression in S. pneumoniae and for identifying virulence genes in S. pneumoniae. Antibodies that selectively bind to the above proteins or peptide fragments can be used to treat S. pneumoniae infection. The antibodies can also be used to detect S. pneumoniae cells.

SQ Sequence 293 AA;

Query Match 0.6%; Score 8; DB 19; Length 293;

CC The protein may be used in a vaccine to prevent or treat  
CC H. pylori infection or to identify H. pylori polypeptide binding  
CC compounds, useful as potential H. pylori life cycle activators or  
CC inhibitors. The genomic sequence of H. pylori (ATCC 55679) was  
CC determined from overlapping contigs generated by mechanically  
CC shearing the bacterial DNA. The sequences were analysed for ORF of  
CC at least 180 nucleotides, and the predicted coding regions defined  
CC by computer evaluation. To identify likely H. pylori antigens for  
CC vaccine development, the amino acid sequences predicted from  
CC various ORF were analysed for significant homology to other known  
CC or exported membrane proteins. Having identified and determined  
CC the sequences of interest, particular regions can be isolated from  
CC H. pylori by PCR amplification for recombinant polypeptide  
CC production, e.g. in E. coli hosts.

XX SQ Sequence 325 AA;

Query Match 0.6%; Score 8; DB 18; Length 325;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 205 NRVGSGAG 212  
Db 153 nrvgsgag 160

RESULT 34  
AAW24644  
ID AAW24644 standard; Protein; 325 AA.

XX AC AAW24644;

XX DT 11-AUG-1997 (first entry)

XX DE H. pylori cytoplasmic protein, 291700.aa.

XX KW Chronic gastritis; duodenal ulcer disease; activator;  
XX KW inhibitor; bacterial life cycle; vaccine; immunisation; detection;  
XX KW antisense; inhibition; cytoplasmic; vacA.

XX OS Helicobacter pylori.

XX FH Key Location/Qualifiers

FT Misc-difference 5 /note= "encoded by CYA"

FT Misc-difference 9 /note= "encoded by ATR"

FT Misc-difference 17 /note= "encoded by GRA"

FT Misc-difference 211 /note= "encoded by AGK"

FT Misc-difference 244 /note= "encoded by AAW"

XX WO9719098-AL

XX PN 29-MAY-1997

XX PD 15-NOV-1996; 96WO-US18542.

XX PF 17-NOV-1995; 95US-0561469.

XX PR (ASTR ) ASTRA AB.

XX PA Smith PH;

XX XX WPI; 1997-298052/27.

XX DR N-PSDB; AAT77462.

XX XX Helicobacter pylori nucleic acid sequences and related proteins -  
XX used for diagnostics and therapeutics

XX PT

Best Local Similarity 100.0%; Pred. No. 28;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1127 ALLQDLNQ 1134

Db 189 allqdlng 196

RESULT 33

AAW20372  
ID AAW20372 standard; protein; 325 AA.

XX AC AAW20372;

XX DT 29-JUL-1997 (first entry)

XX DE H. pylori cytoplasmic protein, 291700.aa.

XX KW Vaccine; prevention; treatment; infection; identification;  
XX KW binding compound; bacterium; life cycle; activator; bacteria;  
XX KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;  
XX KW cytoplasmic.

XX OS Helicobacter pylori.

XX FH Key Location/Qualifiers

FT Misc-difference 5 /label= unknown

FT Misc-difference 9 /note= "encoded by CYA"

FT Misc-difference 17 /label= unknown

FT Misc-difference 211 /note= "encoded by ATR"

FT Misc-difference 234 /label= unknown

FT Misc-difference 240 /note= "encoded by GRA"

FT Misc-difference 244 /note= "encoded by AGK"

FT Misc-difference 247 /note= "encoded by ACS"

FT Misc-difference 253 /note= "encoded by ACS"

FT Misc-difference 253 /label= unknown

FT Misc-difference 253 /note= "encoded by AAW"

FT Misc-difference 253 /note= "encoded by GGS"

FT Misc-difference 253 /note= "encoded by GAR"

XX WO9640893-AL.

XX PN 19-DEC-1996.

XX PD 06-JUN-1996; 96WO-US09122.

XX PF 01-APR-1996; 96US-0630405.

XX PR 07-JUN-1995; 95US-0487032.

XX PA (ASTR ) ASTRA AB.

XX PI Berglindh OT, Smith D, Mellgaerd BL;

XX XX WPI; 1997-052305/05.

XX DR N-PSDB; AAT67782.

XX XX Helicobacter pylori nucleic acid sequences and related  
XX polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
XX infection, and to detect Helicobacter

XX PT Helicobacter pylori cytoplasmic protein.

XX The

20:30

Be  
Helicobacter pylori  
polypeptide(s)  
infection, and  
disclosure, pages 50  
to 53, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000

Claim 18; Pages 166-167; 235pp; English.

XX The present sequence is a Helicobacter pylori cytoplasmic protein, which was found to be homologous to vacA following BLAST protein analysis.

CC H. pylori has been strongly linked to chronic gastritis and duodenal ulcer disease. The nucleic acid sequences of the invention are used to evaluate compounds, especially activators or inhibitors of bacterial life cycle, for the ability to bind an H. pylori nucleic acid sequence. The nucleic acid sequences, and corresponding proteins, are also useful for generating vaccines for immunising subjects against H. pylori or for use in detecting the presence of Helicobacter species in a sample. Antisense nucleic acid sequences of these sequences are used to inhibit expression of a gene from Helicobacter species. H. pylori whole genomic DNA was isolated and nebulised to a median size of 2000 bp. Purified DNA fragments were blunt-ended and ligated to unique BstXI-linker adapters in 100-1000 fold molar excess. These linkers are complementary to the BstXI-cut pMPX vectors. These linkers are not self-complementary. Therefore the linkers will not concatamerise nor will the cut vector re-ligate itself easily. The linker-adaptor inserts were ligated to each of the 20 pMPX vectors to construct a series of shotgun subclone libraries. The purified DNA samples were then sequenced.

CC Note: The ORF/protein reference number for this sequence was obtained from the related specification, WO9640893.

CC Sequence 325 AA;

Query Match 0.68; Score 8; DB 18; Length 325;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
| | | | | | | | | |  
Db 153 nrvgsgag 160

RESULT 35  
AAW55735  
ID AAW55735 standard; Protein; 1213 AA.  
AC AAW55735;  
XX 13-JUL-1998 (first entry)  
DT H. pylori ORF 07es507f9\_960952\_f2\_47 secreted protein.  
DE Cytoplasmic; vaccine prevention; treatment; infection; envelope; identification; binding compound; bacteria; life cycle; activator; inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis; secreted protein.

OS Helicobacter pylori.  
XX WO9737044-A1.  
PN 09-OCT-1997.  
XX 27-MAR-1997; 97WO-US05223.  
XX 06-DEC-1996; 96US-0761318.  
XX 29-MAR-1996; 96US-0625811.  
XX 02-APR-1996; 96US-0758731.  
XX 25-OCT-1996; 96US-0738905.  
XX 28-OCT-1996; 96US-0738859.  
XX (ASTR) ASTRA AB.  
XX Alm RA, Smith D;  
XX WPI; 1997/503122/46.

DR N-PSDB; AAV25144.  
XX Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) - useful in vaccines to treat or prevent H. pylori infection and for diagnosis of H. pylori infection  
PT Disclosure; Page 1007-1010; 1145pp; English.  
XX This sequence represents a Helicobacter pylori secreted protein. The protein may be used in a vaccine to prevent or treat H. pylori infection or to identify H. pylori polypeptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors. The DNA and probes derived from it may be used for the identification of H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic acid sequences complementary to the DNA act as antisense sequences and can be used to prevent the translation of H. pylori mRNA. Antibodies against the protein can be used in immunoassays to evaluate the abundance and distribution of H. pylori-specific antigens. The genomic sequence of H. pylori (ATCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.

CC Sequence 1213 AA;

Query Match 0.68; Score 8; DB 18; Length 1213;  
Best Local Similarity 100.0%; Pred. No. 11e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 NRVGSGAG 212  
| | | | | | | | | |  
Db 185 nrvgsgag 192

RESULT 36  
AAU00023  
ID AAU00023 standard; Protein; 1851 AA.  
AC AAU00023;  
XX 10-MAY-2001 (first entry)  
DT Human activated T-lymphocyte associated sequence 2, ATLAS-2.  
DE Human; activated T-lymphocyte associated sequence 2; ATLAS-2; antibody; cytokine receptor; autoimmune disorder; immune disorder; cancer; T-lymphocyte-associated disorder; cell proliferation disorder; tumour; cell differentiation disorder; immune deficiency disorder; malignancy; viral infection; bacterial infection; fungal infection; metabolism; chromosome 11p15.5.  
XX Homo sapiens.  
OS WO200114564-A2.  
XX 01-MAR-2001.  
XX 18-AUG-2000; 2000WO-US22699.  
XX 20-AUG-1999; 99US-0150105.  
XX 28-APR-2000; 2000US-0560101.  
XX 28-APR-2000; 2000US-0560365.  
XX 28-APR-2000; 2000US-0560948.  
XX 28-APR-2000; 2000US-0561533.  
XX (CURA-) CURAGEN CORP.



✓